

8

SEARCH REQUEST FORM

Scientific and Technical Information Center

Access DB#

Requester's Full Name: _____ Examiner #: _____ Date: _____
Art Unit: _____ Phone Number 30 _____ Serial Number: _____
Mail Box and Bldg/Room Location: _____ Results Format Preferred (circle): PAPER DISK E-MAIL

If more than one search is submitted, please prioritize searches in order of need.

Please provide a detailed statement of the search topic, and describe as specifically as possible the subject matter to be searched. Include the elected species or structures, keywords, synonyms, acronyms, and registry numbers, and combine with the concept or utility of the invention. Define any terms that may have a special meaning. Give examples or relevant citations, authors, etc. if known. Please attach a copy of the cover sheet, pertinent claims, and abstract.

Title of Invention: _____

Inventors (please provide full names): _____

Earliest Priority Filing Date: _____

**For Sequence Searches Only* Please include all pertinent information (parent, child, divisional, or issued patent numbers) along with the appropriate serial number.*

BEST AVAILABLE COPY

STAFF USE ONLY
Contact: Sheppard

Searcher _____

Searcher Phone # tel: 308-4499

Searcher Location _____

Date Searcher Picked Up _____

Date Completed 10/19/01

Searcher Prep & Review Time _____

Clencal Prep Time _____

Online Time _____

Type of Search

NA Sequence (#) _____

AA Sequence (#) _____

Structure (#) _____

Bibliographic _____

Litigation _____

Fulltext _____

Patent Family _____

Other _____

Vendors and cost where applicable

STN _____

Dialog _____

Questel Orbit _____

Dr Link _____

Lexis/Nexis _____

Sequence Systems _____

WWW/Internet _____

Other (specify) _____

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 19:58:14 ; Search time 8321.72 seconds
(without alignments)
3148.785 Million cell updates/sec

Title: US-09-380-337-1
Perfect score: 2772
Sequence: 1 GGTCGCGAGCCGCGACC.....TGTAAAAA 2772

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est25:*
26: gb_est26:*
27: gb_est27:*
28: gb_est28:*
29: gb_est29:*
30: gb_est30:*
31: gb_est31:*
32: gb_est32:*
33: gb_est33:*
34: gb_est34:*
35: gb_est35:*
36: gb_est36:*
37: gb_est37:*
38: gb_est38:*
39: gb_est39:*
40: gb_est40:*
41: gb_est41:*
42: gb_est42:*
43: gb_est43:*
44: gb_est44:*
45: gb_est45:*
46: gb_est46:*
47: gb_est47:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est33:*
111: gb_est34:*
112: gb_est35:*
113: gb_est36:*
114: gb_est37:*
115: gb_est38:*
116: gb_est39:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source

1. .897
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DG002YG12"
/clone_lib="LTI_FL011_BCI"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"
BASE COUNT 152 a 291 c 268 g 182 t
ORIGIN

Query Match 30.5%; Score 844.8; DB 106; Length 897;
Best Local Similarity 98.5%; Pred. No. 7.8e-188;
Matches 879; Conservative 4; Mismatches 6; Indels 3; Gaps 3;

QY 80 GGGCAGGGGCGCGCCACCGCCCGCCGCGCATGGGGCTGAAGCGCCGCCAGAGAGCGT 139
DB 9 GCGCGGAGCGCGCGCCACCGCCCGCCGCGCATGGGGCTGAAGCGCCGCCAGAGAGCGT 68
QY 140 GTTCCCGCTGCGCTCCATCGACGACGCTGGTGGCGCTTTGTGCGCGAGCTGGCGCCGAGA 199
DB 69 GTTCCCGCTGCGCTCCATCGACGACGCTGGTGGCGCTTTGTGCGCGAGCTGGCGCCGAGA 128
QY 200 GGAGCGCGACCTGGTGTCTCTTTCCTTGGTGTGGGCTTCTGGAGACATTTTCTGGGTGT 259
DB 129 GGAGCGCGACCTGGTGTCTCTTTCCTTGGTGTGGGCTTCTGGAGACATTTTCTGGGTGT 188
QY 260 CAACCGCGTATCCCTACCAACGTTCCCGAGCTACCTTCCAGCCCGCCAGCCCGCCCGCA 319
DB 189 CAACCGCGTATCCCTACCAACGTTCCCGAGCTACCTTCCAGCCCGCCAGCCCGCCCGCA 248
QY 320 CCGCGCTGGCGGCTCACCTACTTTCCCGTGGCGACCTGTCTATCATCGCCGCCCTCTA 379
DB 249 CCGCGCTGGCGGCTCACCTACTTTCCCGTGGCGACCTGTCTATCATCGCCGCCCTCTA 308
QY 380 TGCCCGCTTACCGCCCGAGATCCGAGCGCGCTGCGACCTGTCTCTATCTCGAGAAGG 439
DB 309 TGCCCGCTTACCGCCCGAGATCCGAGCGCGCTGCGACCTGTCTCTATCTCGAGAAGG 368
QY 440 GGGTGTCTCCAGCGCTGAGCTGGTGAAGAGGTCTCCGATGTATATGGAACACCTCAG 499
DB 369 GGGTGTCTCCAGCGCTGAGCTGGTGAAGAGGTCTCCGATGTATATGGAACACCTCAG 428
QY 500 CGCGCTCTACTTCAAGATCGGGCCACATCCAGTCCCTCTTCCAGCTTCAATCAGAGCAC 559
DB 429 CGCGCTCTACTTCAAGATCGGGCCACATCCAGTCCCTCTTCCAGCTTCAATCAGAGCAC 488
QY 560 CAAATTTGAGAGCTCCGGTGTGGCCCTTTGCTGTGGTGGGGCTTGCAGGCCCTTGGGTCT 619
DB 489 CAAATTTGAGAGCTCCGGTGTGGCCCTTTGCTGTGGTGGGGCTTGCAGGCCCTTGGGTCT 547
QY 620 CCGGATGTCCACCTCGCCCTGTCTGAGGATCATGCTGGGTAGTGTGGGCGCCATGG 679
DB 548 CCGGATGTCCACCTCGCCCTGTCTGAGGATCATGCTGGGTAGTGTGGGCGCCATGG 607
QY 680 GGAGCAGACAGCTGAGGTACCTTGGCAGCGCAAGGGCAACGAGCGCCAGGGGCCAGAG 739
DB 608 GGAGCAGACAGCTGAGGTACCTTGGCAGCGCAAGGGCAACGAGCGCCAGGGGCCAGAG 667

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	844.8	30.5	897	106	AL560097 AL560097
2	809.2	29.2	912	106	AL582775 AL582775
3	806.2	29.1	876	105	AL521543 AL521543
4	773.6	27.9	865	106	AL560096 AL560096
5	758.4	27.4	808	154	BG476481 BG476481
6	717.7	25.9	718	154	BG471808 BG471808
7	658.2	23.7	898	105	AL521542 AL521542
8	645.2	23.3	799	150	BF526328 BF526328
9	641.4	23.1	966	172	BF980770 BF980770
10	623.8	22.5	959	154	BG476263 BG476263
11	622.2	22.4	981	150	BF525363 BF525363
12	617.2	22.3	709	133	BG396934 BG396934
13	613.4	22.1	680	165	BE267789 BE267789
14	613	22.1	687	172	BF984524 BF984524
15	606	21.9	738	175	BG251389 BG251389
16	601.8	21.7	645	11	AA741057 AA741057
17	598.6	21.6	741	165	BE266998 BE266998
18	592.8	21.4	596	136	BE514777 BE514777
19	566.8	20.4	641	122	AW950638 AW950638
20	566.6	20.4	856	153	BG422364 BG422364
21	560.4	20.2	580	165	BE267618 BE267618
22	554.8	20.0	1139	169	BF795474 BF795474
23	551.6	19.9	764	137	BE573676 BE573676
24	547.6	19.8	898	146	BF237078 BF237078
25	545.4	19.7	547	165	BE274228 BE274228
26	540.6	19.5	630	165	BE281439 BE281439
27	535.4	19.3	687	152	BG326214 BG326214
28	532	19.2	574	122	AW968425 AW968425
29	531.6	19.2	692	143	BF012319 BF012319
30	516.4	18.6	624	114	AW321532 AW321532
31	505.4	18.2	792	106	AL561613 AL561613
32	488.4	17.6	507	112	AW150068 AW150068
33	484	17.5	523	169	BF803247 BF803247
34	477.4	17.2	479	105	AL079746 AL079746
35	466.8	16.8	525	11	AA76738 AA76738
36	466	16.8	533	11	AA743431 AA743431
37	461.8	16.7	525	19	AI333944 AI333944
38	460	16.6	532	136	BE532045 BE532045
39	459	16.6	503	15	AI082242 AI082242
40	449.4	16.2	497	104	AI953167 AI953167
41	448.8	16.2	496	116	AW439193 AW439193
42	445	16.1	490	152	BG328937 BG328937
43	444	16.0	511	10	AA703195 AA703195
44	443.8	16.0	635	136	BE533157 BE533157
45	442.8	16.0	457	191	Z98466 HS298466

ALIGNMENTS

RESULT 1
AL560097
LOCUS AL560097 LTI_FL011_BCI Homo sapiens cDNA clone CS0DG002YG12 5 prime
DEFINITION AL560097 897 bp mRNA EST 16-FEB-2001
AL560097 LTI_FL011_BCI Homo sapiens cDNA clone CS0DG002YG12 5 prime
ACCESSION AL560097
VERSION 1 (bases 1 to 897)
KEYWORDS mRNA sequence.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 897)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope

BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB003YA04"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/lab_host="DH10B"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 153 a 281 c 256 g 180 t 6 others
ORIGIN

Query Match 29.1%; Score 806.2; DB 105; Length 876;
Best Local Similarity 96.4%; Pred. No. 9,1e-179;
Matches 845; Conservative 6; Mismatches 10; Indels 16; Gaps 2;

QY 90 CGCGGCCACCGCCGCCATGAGGGCTGAAGCGGCCAGAACGCTGTTCGGCTG 149
Db 1 CGCGGCCACCGCCGCCATGAGGGCTGAAGCGGCCAGAACGCTGTTCGGCTG 60
QY 150 CGCTCATCGACGCTGTGGCGCTGTTGCTGCCAGCTGGCGGAGGAGCGCGAC 209
Db 61 CGCTCATCGACGCTGTGGCGCTGTTGCTGCCAGCTGGCGGAGGAGCGCGAC 120
QY 210 CTGGTGCTCTTCCCTTGGTGTGGCTGCTGAGCAATTTCTGGCTGTCAACCGCTC 269
Db 121 CTGGTGCTCTTCCCTTGGTGTGGCTGCTGAGCAATTTCTGGCTGTCAACCGCTC 180
QY 270 ATCCCTACCAAGTTCGCGAGCTCACCCTCCAGCCCGCCGCCCGCGCGCTGCG 329
Db 181 ATCCCTACCAAGTTCGCGAGCTCACCCTCCAGCCCGCCGCCCGCGCGCTGCG 240
QY 330 GGCCTCACCTACTTCCCGTGGCGACCTGTCTATCATCGCGCCCTCTATGCCGCTTC 389
Db 241 GGCCTCACCTACTTCCCGTGGCGACCTGTCTATCATCGCGCCCTCTATGCCGCTTC 300
QY 390 ACCGCCAGATCCGAGGCGCCGTCGACCTGTCCCTCTATCTCGAAGGGGTGTCTCC 449
Db 301 ACCGCCAGATCCGAGGCGCCGTCGACCTGTCCCTCTATCTCGAAGGGGTGTCTCC 360
QY 450 AGCGGTGAGCTGGTGAAGAGTCTCCGATGTCATGGAACAGCTCAGCGCTCTCTAC 509
Db 361 AGCGGTGAGCTGGTGAAGAGTCTCCGATGTCATGGAACAGCTCAGCGCTCTCTAC 420
QY 510 TTCAAGGATCGGGCCACATCCAGTCCCTCTCAGCTTCATCAC-----A 554
Db 421 TTCAAGGATCGGGCCCAATCCAGTCCCTCTCAGCTTCATCACAGTTGGAGCCCAAGTA 480
QY 555 GGCACCAAAATGGACAGCTCCGCTGTGGCTTTGCTGTGGTGGGCCCTGCCAGGCCCTG 614
Db 481 GGCACCAAAATGGACAGCTCCGCTGTGGCTTTGCTGTGGTGGGCCCTGCCAGGCCCTG 539
QY 615 GGTCTCCGGATGTCCACCTCGCCCTGTCTGAGGATCATGCGCTGGGTAGTGTGGGCC 674
Db 540 GGTCTCCGGATGTCCACCTCGCCCTGTATGAGGATCATGCGCTGGGTAGTGTGGGCC 599
QY 675 AATGGGAGCAGACGCTGAGTCACTGGCAGCGGAAGGGCAACAGGACCGCAGGGC 734
Db 600 AATGGGAGCAGACGATGAGTCACTGGCAGCGGAAGGGCAACAGGACCGCAGGGC 659

QY 735 CAGACAGTCAATCCGCTGCTGGCTGAGCGAGCTGCTGTACCTGAAGGATCATACATG 794
Db 660 CAGACAGTCAATCCGCTGCTGGCTGAGCGAGTGGBCCTGTACTCTGAAGGATCATACATG 719
QY 795 CGCTGTGACCGCAAGATGAGGTGGCTTCATGGTGTGCCATCAACCTTCCATTGAC 854
Db 720 CGCTGTGACCGCAAGATGAGGTGGCTTCATGGTGTGCCATCAACCTTCCATTGAC 779
QY 855 CTGCACACGACTCGCTGGAGCTTCTGCAGCTGCACGAGAAGCTGCTCTGGCTGCTCTAT 914
Db 780 CTGCACACGACTCGCTGGAGCTTCTGCAGCTGCACGAGAAGCTGCTCTGGCTGCTCTAT 839
QY 915 GACCTGGGACATCTGGAAAGGTACCCCATGGCCTTAG 951
Db 840 GACCTGGGACATCTGGAAAGGTACCCCATGGCCTTAG 876

RESULT 4
AL560096/c 865 bp mRNA EST 16-FEB-2001
LOCUS AL560096 LTI_FL011_BC1 Homo sapiens cDNA clone CS0DB02YG12 3 prime
DEFINITION mRNA sequence.
ACCESSION AL560096
VERSION AL560096.1 GI:12906228
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS Li.W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
Location/Qualifiers
1. 865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DB002YG12"
/clone_lib="LTI_FL011_BC1"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 165 a 183 c 306 g 204 t 7 others
ORIGIN

Query Match 27.9%; Score 773.6; DB 106; Length 865;
Best Local Similarity 97.1%; Pred. No. 4,1e-171;
Matches 845; Conservative 5; Mismatches 14; Indels 6; Gaps 6;

QY 1700 AGGTGGCAGCAGGCTCAGTGGCAGCACCCG-CAGCATCACCCAGCCGAGGGTCCAG 1758
Db 865 AAGTGGCAGCAGGCTCAGTGGCAGCACCCGAGCATCACCCAGCCGAGGGTCCAG 806
QY 1759 TGCTCACTTTCCAGAGTGAAGATGAAGGCATGAAGGAGCTGCTGTCGCCACCAAGA 1818
Db 805 TRCTCACTTTCCAGAGTGAAGATGAAGGCATGAAGGAGCTGCT-GTGGCCACCAAGA 747
QY 1819 TCAACTCGAGCGCCATCAAGCTGCAACTCAGGCACAGTCCGACAGTGCAGATGAAGAAGC 1878

Db 746 TCAACTCGAGCGCCATCAAGTGCACACTCAGGCACAGTGCAGATGAAGAAGC 687
Qy 1879 AGAAGTGTCCACCCCTAGTACACTCTGTCTTCTCAAGCGGACGCGGAAGGCC 1938
Db 686 AGAAGTGTCCACCCCTAGTACACTCTGTCTTCTCAAGCGGACGCGGAAGGCC 627
Qy 1939 TCTGAACCTACTGGGACCTTCGGACCGCTTGTGGGACCCAGGCTCCGCTTAGTCCGCCA 1998
Db 526 TCTGAACCTACTGGGACCTTCGGACCGCTTGTGGGACCCAGGCTCCGCTTAGTCCGCCA 567
Qy 1999 ACTCTAGAGCCATGTTCTGCCCCAGCCCCAAAGGGGACAGGCTTCACTCTACCCAAACC 2058
Db 566 ACTCTAGAGCCATGTTCTGCCCCAGCCCCAAAGGGGACAGGCTTCACTCTACCCAAACC 507
Qy 2059 CTAGGTTCCGGTCCCGAGTACAGTCTGTATCAACCCAGGATTTTCTCCAGCTCAGAAC 2118
Db 506 CTAGGTTCCGGTCCCGAGTACAGTCTGTATCAACCCAGGATTTTCTCCAGCTCAGAAC 447
Qy 2119 CCAGGGCTCTGCCCCAGCTGTTAGATATAGTCTTCTTCCAGAAATCCAGCGGCCA 2178
Db 446 CCA-GGCTCTGCCCCAGCTGTTAGATATAGTCTTCTTCCAGAAATCCAGCGGCCA 388
Qy 2179 ATGGAACCTACAGCTGGTCTTAATACAGTCTTTAAAGGCCAGCCGCTAGAAACCC 2238
Db 387 ATGGAACCTACAGCTGGTCTTAATACAGTCTTTAAAGGCCAGCCGCTAGAAACCC 328
Qy 2239 AAGCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTCTTACTCAGGGCTCTCCCA 2298
Db 327 AAGCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTCTTACTCAGGGCTCTCCCA 268
Qy 2299 GCTTGTAGGAGCTGAGTCTTACCCCTTAACCCAAAGGGGACAGGTCCTCCACCTCCAGCCC 2358
Db 267 GCTTGTAGGAGCTGAGTCTTACCCCTTAACCCAA-GGAGCACAGTCTCCACCMCMCCC 209
Qy 2359 GGGAGCTTAGGACCTACAGCCCTAGGAGTATATTTCGCACCTTCAGAAATCCCATATC 2418
Db 208 -GGAGCTTAGGACCTACAGCCCTAGGAGTATATTTCGCACCTTCAGAAATCCCATATC 150
Qy 2419 TTGGAAATCAAGCTCCGCTGCCCCAAATAACTTCAGTCTTCCAGAAATTTGGAATC 2478
Db 149 TTGGAAATCAAGCTCCGCTGCCCCAAATAACTTCAGTCTTCCAGAAATTTGGAATC 91
Qy 2479 CTAGTTTCTCTCTCTGATATCCGAGTCTGGGACACAAACTCCGCCCCCAGGCTATGA 2538
Db 90 CTAGTTTCTCTCTCTGATATCCGAGTCTGGGACACAAACTCCGCCCCCAGGCTATGA 31
Qy 2539 GCATCTGAGCCCGCCCTCTTCCCTGACGA 2568
Db 30 GCATCTGAGCCCGCCCTCTTCCCTGACGA 1

RESULT 5
BG476481 808 bp mRNA EST 21-MAR-2001
LOCUS 60252015F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4640450 5',
DEFINITION mRNA sequence.
ACCESSION BG476481
VERSION BG476481.1 GI:13408760
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 808)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC).
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTD/DRP
cDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Plate: LLC1405 row: n column: 03
High quality sequence stop: 798.
Location/Qualifiers
source
1. .808
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4640450"
/tissue_type="melanotic melanoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: skin; Vector: pOTB7; Site_1: xhoi; Site_2: EcoRI; cDNA made by oligo-dT priming. Directionally cloned into EcoRI/XhoI sites using the following 5' adaptor: GGCACGAG(G). Size-selected >500bp for average insert size 1.8kb. Library constructed by Ling Hong in the laboratory of Gerald M. Rubin (University of California, Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT 177 a 245 c 278 g 108 t
ORIGIN
Query Match 27.4%; Score 758.4; DB 154; Length 808;
Best Local Similarity 99.4%; Pred. No. 1.5e-167;
Matches 803; Conservative 0; Mismatches 1; Indels 4; Gaps 4;
Qy 1092 CACTGTGCGAACCGCAATGTGCGGGAAGCCCTGCAGCGCTGGGCGGACACGGCCACTGTC 1151
Db 2 CACTGTGCGAACCGCAATGTGCGGGAAGCCCTGCAGCGCTGGGCGGACACGGCCACTGTC 61
Qy 1152 ATCCAGGACTACAACTACTGCCGGAAGACGAGGAGATCTACAAGGAGTCTTTTGAAGTA 1211
Db 62 ATCCAGGACTACAACTACTGCCGGAAGACGAGGAGATCTACAAGGAGTCTTTTGAAGTA 121
Qy 1212 GCCAATGATGTATCCCAACCTGCTGGAAGGAGGACGACCACTTGTGTGAGGCGGCGGAG 1271
Db 122 GCCAATGATGTATCCCAACCTGCTGGAAGGAGGACGACCACTTGTGTGAGGCGGCGGAG 181
Qy 1272 GAGGCGCGGGGAGCAAGCCAGGCGCACCCAGAGCCAAAGGTTCCGCGCTCCAGGACCCCT 1331
Db 182 GAGGCGCGGGGAGCAAGCCAGGCGCACCCAGAGCCAAAGGTTCCGCGCTCCAGGACCCCT 241
Qy 1332 GAGTGTCTTCCGCCACCTGCTGCGATTCTACGACGCGATCTGCAATGGAGAGGCGCAGT 1391
Db 242 GAGTGTCTTCCGCCACCTGCTGCGATTCTACGACGCGATCTGCAATGGAGAGGCGCAGT 301
Qy 1392 CCCACGCTGTGTGACAGTGGGCTGGGCGCACCTTTCTTGTGACAGTCCCTAGGCCGCTTTT 1451
Db 302 CCCACGCTGTGTGACAGTGGGCTGGGCGCACCTTTCTTGTGACAGTCCCTAGGCCGCTTTT 361
Qy 1452 GAGGACAGGTGCGGCAAGAGTTCGCATAGTACGCGGAGGCGGAGGCGGCGGCGGCGG 1511
Db 362 GAGGACAGGTGCGGCAAGAGTTCGCATAGTACGCGGAGGCGGAGGCGGCGGCGGCGG 421
Qy 1512 GAGGAGCGGTGGGCGGAGGAGCCCGGGAAGCGCGGCGGCGGCGGCGGCGGCGGAGTCC 1571
Db 422 GAGGAGCGGTGGGCGGAGGAGCCCGGGAAGCGCGGCGGCGGCGGCGGCGGCGGAGTCC 481
Qy 1572 AAGCCAGAGAGCCCGCGCCCAAGAGGACGACCTGGACAGGCGCTGGGCGGCGGCGG 1631
Db 482 AAGCCAGAGAGCCCGCGCCCAAGAGGACGACCTGGACAGGCGCTGGGCGGCGGCGG 541
Qy 1632 CAGGTCAGTGTGAGGACCCCGCGGGAAGCTCTCGGAGCTGTCGCTGGCACAGCCGGA 1691
Db 542 CA-GGTGAGTGTGAGGACCCCGCGGGAAGCTCTCGGAGCTGTCGCTGGCACAGCCGGA 600
Qy 1692 GGGCCTGAAGTGTGAGCAGCGGCTCAGGTGCGCAGACCCCGGAGCATCACACCCCGGAG 1751
Db 601 GGGCCTGAAGTGTGAGCAGCGGCTCAGGTGCGCAGACCCCGGAGCATCACACCCCGGAG 660

Query Match	23.3%	Score	645.2	DB	150	Length	799
Best Local Similarity	95.2%	Pred. No.	6e-141				
Matches	687	Conservative	0	Mismatches	33	Indels	2
Gaps	2						
Qy	1300	CCAGAGCCAAGTTCCGCCCTCCAGGACCCTGAGTGGTTCCGCCACCTGCTGCGATTCT	1359				
Db	1	CCCAGAGCCAAGTTCCGCCCTCCAGGACCCTGAGTGGTTCCGCCACCTGCTGCGATTCT	60				
Qy	1360	ACGAGCGCATCTGCAATATGGGAGGAGGSGACTGCCACGCTGTGTCACCTGGGCTGGG	1419				
Db	61	ACGAGCGGCATCTGCAATATGGGAGGAGGSGACTGCCACGCTGTGTCGACACTGGGCTGGG	120				
Qy	1420	CCACCTTTCTTTGTGCAGTCCCTAGGCCCTTTTGGGGACAGGTGGGCGAGAAGTGGCGCA	1479				
Db	121	CCACCTTTCTTTGTGCAGTCCCTAGGCCCTTTTGGGGACAGGTGGGCGAGAAGTGGCGCA	180				
Qy	1480	TAGTGAGCCGAGAGGCCGAGCGCCGAGGCGCGTAGGACCTTGGGGCGAGGAAGCCCGGG	1539				
Db	181	TAGTGAGCCGAGAGGCCGAGCGCCGAGGCGCGTAGGACCTTGGGGCGAGGAAGCCCGGG	240				
Qy	1540	AAGCGCGCGGGGGCCCGAGCGGGAGTCCAGCCAGAGAGAGGCCCGCGGCCCAAGA	1599				
Db	241	AAGCGCGCGGGGGCCCGAGCGGGAGTCCAGCCAGAGAGAGGCCCGCGGCCCAAGA	300				

QY 1600 AGCCAGCACTGGACAAGGGCCCTGGGACCGCCGAGGTGCAGTGTTCAGGACCCCGCCGGA 1659
|||||
Db 301 AGCCAGCACTGGACAAGGGCCCTGGGACCGCCGAGGTGCAGTGTTCAGGACCCCGCCGGA 360
|||||
QY 1660 AGCCTCTGGGACTGTCGCTGGGCACACCGCCGAGGCTGAAGTGGCAGCAGCGCTCAGG 1719
|||||
Db 361 AGCCTCTGGGACTGTCGCTGGGCACACCGCCGAGGCTGAAGTGGCAGCAGCGCTCAGG 420
|||||
QY 1720 TGGCAGCAGCCCGCAGCATCACCAACCGCCGAGGCTCAGTGCCTTCCTCCAGAGTGAGA 1779
|||||
Db 421 TGGCAGCAGCCCGCAGCATCACCAACCGCCGAGGCTCAGTGCCTTCCTCCAGAGTGAGA 480
|||||
QY 1780 AGATGAAGGCGATGAAGAGCTGCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGC 1839
|||||
Db 481 AGATGAAGGCGATGAAGAGCTGCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGC 540
|||||
QY 1840 TGCAACTCAGGCACAGTCGCAAGTGCAGATGAAGAG - CAGAAAGTGTCCACCCCTAGT 1898
|||||
Db 541 TGCAACTCAGGCACAGTCGCAAGTGCAGATGAAGAG - CAGAAAGTGTCCACCCCTAGT 600
|||||
QY 1899 GACTACACTGTCTCTTCTCAAGCGGCGCAGCGCAAGGCTCTGAACTACTTGGGGACTTC 1958
|||||
Db 601 GACTACACTGTCTCTTCTCAAGCGGCGCAGCGCAAGGCTCTGAACTACTTGGGGACTTC 660
|||||
QY 1959 GGACCGCTTGTGGGGACCCAGGCTCGCCCTTAGTCCCGCAACTCTGAGCCCATGTTCTGC 2018
|||||
Db 661 GGACCGCTTGTGGGGA - CCAGGATCGCAGTAAAGTCCCGCAACTCTGAGCCCATGTTCTGC 719
|||||
QY 2019 CC 2020
||
Db 720 CC 721

RESULT 9

BF980770 966 bp mRNA EST 23-JAN-2001
LOCUS 602303938F1 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:4395417 5',
mRNA sequence.
ACCESSION BF980770
VERSION BF980770.1 GI:12383582
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE NIH-MGC <http://mgc.nci.nih.gov/>.
AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE Unpublished (1999)
JOURNAL Contact: Robert Strausberg, Ph.D.
COMMENT Email: cgapbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM10092 row: 1 column: 10
High quality sequence start: 3
High quality sequence stop: 692.
Location/Qualifiers

FEATURES
source

1. .966
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4395417"
/clone_lib="NIH_MGC_88"
/tissue_type="duodenal adenocarcinoma, cell line"
/lab_host="PH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site_1: NotI; Site_2: SalI; Cloned unidirectionally;
oligo-dr primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life

BASE COUNT 208 a 280 c 301 g 177 t
ORIGIN
Query Match 23.1%; Score 641.4; DB 172; Length 966;
Best Local Similarity 96.0%; Pred. No. 4.9e-140;
Matches 692; Conservative 0; Mismatches 21; Indels 8; Gaps 3;
QY 349 TGGCCGACCTGTATCATCGCGCCCTCTATGCCGCTTACCGCCACATCCGAGCG 408
|||||
Db 1 TGGCCGACCTGTATCATCGCGC - CCTCTATGCCGCTTACCGCCACATCCGAGCG 59
|||||
QY 409 CCCTCGACCTGTCCCTCTATCTTCGAGAGGGGTGTCTCCAGCCGTGAGTGGTGAAGA 468
|||||
Db 60 CCCTCGACCTGTCCCTCTATCTTCGAGAGGGGTGTCTCCAGCCGTGAGTGGTGAAGA 119
|||||
QY 469 AGTCTCCGATGTCATATGGAACAGCCTTCAGCCGCTCTACTTCAAGGATCGGCCCCACA 528
|||||
Db 120 AGTCTCCGATGTCATATGGAACAGCCTTCAGCCGCTCTACTTCAAGGATCGGCCCCACA 179
|||||
QY 529 TCAGTCCCTCTTCAGCTTCATCACAGGACCAANTTGGACAGCTCCGGTGTGGCCTTTG 588
|||||
Db 180 TCAGTCCCTCTTCAGCTTCATCACAGGACCAANTTGGACAGCTCCGGTGTGGCCTTTG 239
|||||
QY 589 CTGTGTGGGGCTTGCAGGCGCTTGGGTCTCGGGATGTCCACCTCGCCCTGTCTCAGG 648
|||||
Db 240 CTGTGTGGGGCTTGCAGGCGCTTGGGTCTCGGGATGTCCACCTCGCCCTGTCTCAGG 299
|||||
QY 649 ATCATGCCCTGGGTAGTGTTTGGGCCCAATGGGAGCAGACAGCTGAGGTCACTTGGCACG 708
|||||
Db 300 ATCATGCCCTGGGTAGTGTTTGGGCCCAATGGGAGCAGACAGCTGAGGTCACTTGGCACG 359
|||||
QY 709 GCAAGGCAACGAGGACCGCCAGGCGCCAGACAGTCAATGCCGTGTGGCTGAGCGGAGCT 768
|||||
Db 360 GCAAGGCAACGAGGACCGCCAGGCGCCAGACAGTCAATGCCGTGTGGCTGAGCGGAGCT 419
|||||
QY 769 GCGTGTACCTGAAAGGATCATATGCGCTGTGACCCCAAGATGGAGGTGCGCTTCATGG 828
|||||
Db 420 GCGTGTACCTGAAAGGATCATATGCGCTGTGACCCCAAGATGGAGGTGCGCTTCATGG 479
|||||
QY 829 TGTGTGCCATCAACCCCTTCATTGACCTGCACACCGCACTCGCTGGAGCTTCTGCAGTGC 888
|||||
Db 480 TGTGTGCCATCAACCCCTTCATTGACCTGCACACCGCACTCGCTGGAGCTTCTGCAGTGC 539
|||||
QY 889 ACAGAGAGCTGTCTGGCTCTATGACCTTGGACATCTGGAAGAGTACCCCATGSCCT 948
|||||
Db 540 ACAGAGAGCTGTCTGGCTCTATGACCTTGGACATCTGGAAGAGTACCCCATGSCCT 599
|||||
QY 949 TAGGGAACCTGGCAGATCTAGAGGAGCTGGAGCCGCCCTTGGCCGCGCAGACCCACTCA 1008
|||||
Db 600 TAGGGAACCTGGCAGATCTAGAGGAGCTGGAGCCGCCCTTGGCCGCGCAGACCCACTCA 657
|||||
QY 1009 CCCTCTACCAACAAGGCGATTGCTCCCTAGCCCAAGACCTACTATCGGGATGAACATCTACC 1068
|||||
Db 658 CCTCTTACCAAGGCGATGCT - - - - - CAGCAAGACCTACTATCGGGATGAACACCTTTAGC 712
|||||
QY 1069 C 1069
||
Db 713 C 713

RESULT 10

BG476263 959 bp mRNA EST 21-MAR-2001
LOCUS 602525128F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4643383 5',
mRNA sequence.
ACCESSION BG476263
VERSION BG476263.1 GI:13408542
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 959)
NIH-MGC http://mgc.nci.nih.gov/

National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: ATCC/DCTD/DTF

cDNA Library Preparation: Ling Hong/Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLC1413 row: h column: 08

High quality sequence stop: 715.

Location/Qualifiers

1..959

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4643383"

/clone_lib="NIH_MGC_20"

/tissue_type="melanotic melanoma"

/lab_host="DH10B (phage-resistant)"

/note="Organ: skin; Vector: pOTB7; Site_1: XhoI; Site_2:

EcoRI; cDNA made by oligo-dT priming. Directionally

cloned into EcoRI/XhoI sites using the following 5'

adaptor: GGCACGAG(G). Size-selected >500bp for average

insert size 1.8kb. Library constructed by Ling Hong in

the laboratory of Gerald M. Rubin (University of

California, Berkeley) using ZAP-cDNA synthesis kit

(Stratagene) and Superscript II RT (Life Technologies)."

214 a 281 c 322 g 142 t

BASE COUNT

ORIGIN

Query Match 22.5%; Score 623.8; DB 154; Length 959;

Best Local Similarity 91.7%; Pred. No. 6.6e-136;

Matches 826; Conservative 0; Mismatches 57; Indels 18; Gaps 15;

Qy 1092 CACTGTCCCAACCGCAATGTGGGGAAGCCCTGCGAGGCTGCGGGGACACGCCCACTGTC 1151

Db 2 CACTGTCCCAACCGCAATGTGGGGAAGCCCTGCGAGGCTGCGGGGACACGCCCACTGTC 61

Qy 1152 ATCCAGGACTCACTACTGCGGGGAAGACGAGGAGATCTACAAGGAGTCTTTGAAGTA 1211

Db 62 ATCCAGGACTCACTACTGCGGGGAAGACGAGGAGATCTACAAGGAGTCTTTGAAGTA 121

Qy 1212 GCCAATGATGTCATCCCAACCTGCTGAAGGAGGAGCCAGCTTGTGAGGCGGGCGAG 1271

Db 122 GCCAATGATGTCATCCCAACCTGCTGAAGGAGGAGCCAGCTTGTGAGGCGGGCGAG 181

Qy 1272 GAGCGCGGGGAGCAAGCCAGGCGACCCAGAGCCAGCTTCGCCCTCCAGGACCCCT 1331

Db 182 GAGCGCGGGGAGCAAGCCAGGCGACCCAGAGCTTCGCCCTCCAGGACCCCT 240

Qy 1332 GAGTCTTTCGCCCACTGCTGCGATTCTACGACGCACTCTGCAATGGAGGAGGCG-AG 1390

Db 241 GAGTCTTTCGCCCA-CTGCTGCGATTCTACGACGCACTCTGCAATGGAGGAGGCGCAAG 299

Qy 1391 TCCCAAGCGCTGTGTCAGTGGGTGGGCGACCTTCTTGTGCACTCCCTAGGCGGTTT 1450

Db 300 TCCCAAGCGCTGTGTCAGTGGGTGGGCGACCTTCTTGTGCACTCCCTAGGCGGTTT 358

Qy 1451 TGAGGAGAGTGGCGGAGAGTGGCGATAGTAGCGAGGAGGCGGCGGCGGAGGC 1510

Db 359 TGAGGAGAGTGGCGGAGAGTGGCGATAGTAGCGAGGAGGCGGCGGCGGAGGC 417

Qy 1511 CGAGGAGCGGTGGGCGAGGAGCGCGGAGGCGGCGGCGGCGGCGGCGGAGTC 1570

Db 418 CGAGGAGCGGTGGGCGAGGAGCGCGGAGGCGGCGGCGGCGGCGGCGGAGTC 477

Qy 1571 CAAGCCAGGAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1630

Db 478 CAAGCCAGAGAGGCCCGCGCGCAAGAACGACACTGGAACAGGCGCTTGGCACCGG 537

Qy 1631 CCAGGGTGCAGTGTTCAGGAGCCCCCCC-GGAAGCCTCTCTGGG-ACTGTGCTGGCACAGCC 1688

Db 538 CCAGGGTGCAGTGTTCAGGAGCCCCCCC-GGAAGCCTCTCTGGG-ACTGTGCTGGCACAGCC 597

Qy 1689 CGAGGGCCTTGAAGTGGCAGCAGCGGTTCAGGTGGCAGCAGCCCG-AGCATCACCA--CCG 1745

Db 598 CGAGGGCCTTGAAGTGGCAGCAGCGGTTCAGGTGGCAGCAGCCCG-AGCATCACCA--CCG 657

Qy 1746 CCGGAGGGTCCAGT-GCTCACTTTCAGAGT-GAGAGAGTCAAGGCGATGAAGGAGTGC 1803

Db 658 CCGGAGGGTCCAGTGTTCAGAGTGTTCAGAGTGGAGAGTGCAGGCGATGAAGGAGTGA 717

Qy 1804 TGTGGCCACCAAGATCAACTCGAGCGCCAT--CAAGCTGCAACTCAGGCGACAGTGC 1861

Db 718 TGTGGCCACCAAGATC-ACTCGAGCGCCATTCGAAGCCTGCAGTGCAGGCGACAGTGC 776

Qy 1862 AGTCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTGTACTACTCTGTCTTCTCTCAA 1921

Db 777 GTTCAGATGAGGAAGCAGAAAGTGTCCACCCCTAGTGT--TTAACTCTGTCTCTCAA 834

Qy 1922 CGCGCAGCGCAAGGCGCTC-TGAACCTACTGGGAGCTTCGGAGCCGCTTGTGGGAGCCAGG 1980

Db 835 CGCGCAGCGCAAGGCGCTTCTTGAACCTAATGGGAGTACAGGACCCGTTGTGGAGCCAGGT 894

Qy 1981 C 1981

Db 895 C 895

RESULT 11

BF525363 981 bp mRNA EST 11-DEC-2000

LOCUS 602069440F1 NCLCGAP_Brn64 Homo sapiens cDNA clone IMAGE:4212704

DEFINITION S, mRNA sequence.

ACCESSION BF525363

VERSION BF525363.1 GI:11612724

KEYWORDS EST.

SOURCE human.

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.

1 (bases 1 to 981)

NIH-MGC http://mgc.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: David N. Louis, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLAM9783 row: o column: 09

High quality sequence stop: 730.

Location/Qualifiers

1..981

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:4212704"

/clone_lib="NCLCGAP_Brn64"

/tissue_type="glioblastoma with EGFR amplification"

/lab_host="DH10B (T1 phage-resistant)"

/note="Organ: brain; Vector: pCMV-SPORT6; Site_1: NotI;

Site_2: SalI; Cloned unidirectionally. Primer: pCMV-SPORT6; Oligo dr.

Average insert size 1.57 kb. Constructed by Life

Technologies. Note: this is a NCI_CGAP library."

231 a 290 c 326 g 134 t

BASE COUNT

ORIGIN

FEATURES

source

Query Match 22.4%; **Score** 622.2; **DB** 150; **Length** 981;
Best Local Similarity 91.0%; **Pred. No.** 1.6e-135;
Mismatches 698; **Conservative** 0; **Indels** 11; **Gaps** 3;

QY	857	GCACACCGACTCGCTGGAGCTTCTGCAGCTGCAGACAGAGCTGCTGCTGCTCTATGA	916
DB	2	GCACACCGACTCGCTGGAGCTTCTGCAGCTGCAGACAGAGCTGCTGCTGCTCTATGA	61
QY	917	CTTGGGACATCTGGAAGGTACCCATGCGCTTATGGGAACCTGGCAGATCTAGAGGACT	976
DB	62	CTTGGGACATCTGGAAGGTACCCATGCGCTTATGGGAACCTGGCAGATCTAGAGGACT	121
QY	977	GGAGCCCAACCCCTGGCGGCGACACCACTACCCCTTACCAAGAGGCAATGGCTCAGC	1036
DB	122	GGAGCCCAACCCCTGGCGGCGACACCACTACCCCTTACCAAGAGGCAATGGCTCAGC	181
QY	1037	CAAGACCTACTATCGGGATGACACATCTACCCCTACATGTACCTGGCTGCTACCACTG	1096
DB	182	CAAGACCTACTATCGGGATGACACATCTACCCCTACATGTACCTGGCTGCTACCACTG	241
QY	1097	TCGCAACCGCAATGTGCGGGAAGCCCTGCGAGCCCTGGCGGACGACGCCACTGTATCCA	1156
DB	242	TCGCAACCGCAATGTGCGGGAAGCCCTGCGAGCCCTGGCGGACGACGCCACTGTATCCA	301
QY	1157	GGACTACAATCTACTGCGGGAAGACGAGGAGATCTACAAGAGGTCTTTTGAAGTAGCCAA	1216
DB	302	GGACTACAATCTACTGCGGGAAGACGAGGAGATCTACAAGAGGTCTTTTGAAGTAGCCAA	361
QY	1217	TGATGTATCCCAACCTGCTGAAGGAGCGACCGAGCTTGTGTGAGGCGGGCGAGGAGCG	1276
DB	362	TGATGTATCCCAACCTGCTGAAGGAGCGACCGAGCTTGTGTGAGGCGGGCGAGGAGCG	421
QY	1277	GCGGGGAGCAAGACGAGGACCAAGCAAGGTTCCGCGCTCCAGGACCCCTGAGTG	1336
DB	422	GCGGGGAGCAAGACGAGGACCAAGGTTCCGCGCTCCAGGACCCCTGAGTG	480
QY	1337	CTTTCGCGCCACTGCTGCGATTTCTACACGCGCATCTGCAAAATGGGAGGCGGAGTCCAC	1396
DB	481	CTTTCGCGCCACTGCTGCGATTTCTACACGCGCATCTGCAAAATGGGAGGCGGAGTCCAC	540
QY	1397	GCTGTGTGCTGACGTGGCTGGCGCCACCTTTCTGTGACGTCCCTAGGCGGCTTTTGAAGG	1456
DB	541	GCTGTGTGCTGACGTGGCTGGCGCCACCTTTCTGTGACGTCCCTAGGCGGCTTTTGAAGG	600
QY	1457	ACAGGTGCGGAGAGGTGGCATAGTACGCGAGAGCGGCGGAGGCGGCGGAGGCGGAGGA	1516
DB	601	ACAGGTGCGGAGAGGTGGCATAGTACGCGAGAGCGGCGGAGGCGGAGGCGGAGGAGGA	658
QY	1517	GCGGTGGGCGAGGAGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGAGTCCAGGCC	1576
DB	659	ACGTGCGGCGAGAGCGCGG-----AAGCGGGGCGGCGGCGGCGGAGTCAAGCC	710
QY	1577	AGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	1623
DB	711	AGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG	757

RESULT 12
BG396934
LOCUS 602433842f1 NTH_MGC_20 Homo sapiens cDNA clone IMAGE:4551874 5',
DEFINITION mRNA sequence.
ACCESSION BG396934
VERSION BG396934.1 GI:13290382
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DTP
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LCM1247 row: c column: 11
 High quality sequence stop: 699.

FEATURES
 source
 1..709
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:4551874"
 /clone_lib="NIH_MGC_20"
 /tissue_type="melanotic melanoma"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: skin; Vector: pORF7; Site:1. XhoI; Site:2.
 EcoRI; cDNA made by oligo-dT priming. Directionally
 cloned into EcoRI/XhoI sites using the following 5'
 adaptor: GGCACGAG(G). Size-selected ~500bp for average
 insert size 1.8kb. Library constructed by Ling Hong in
 the laboratory of Gerald M. Rubin (University of
 California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies)."
 BASE COUNT 151 a 214 c 253 g 91 t
 ORIGIN
 Query Match 22.3%; Score 617.2; DB 153; Length 709;
 Best Local Similarity 97.3%; Pred. No. 2.2e-134;
 Matches 691; Conservative 0; Mismatches 13; Indels 6; Gaps 6
 QY 1092 CACTGTGCGCAACCGCAATGTGCGGGAA-GCCCTGCGAGGCTTGGCGGACACGCCACTGT 1150
 DB 2 CACTGTGCGCAACCGCAATGTGCGGGAAAGCCCTGCGAGGCTTGGCGGACACGCCACTGT 61
 QY 1151 CATCCAGGACTACAACTACTGCGGGGAAGACGAGGAGATCTACAAGAGTTCTTTGAAGT 1210
 DB 62 CATCCAGGACTACAACTACTGCGGGGAAGACGAGGAGATCTACAAGAGTTCTTTGAAGT 121
 QY 1211 AGCCAAATGATGTATCCCCAACCTGCTGAGGAGGAGGAGGAGTCTGCTGGAGCGGGCGA 1270
 DB 122 AGCCAAATGATGTATCCCCAACCTGCTGAGGAGGAGGAGGAGTCTGCTGGAGCGGGCGA 181
 QY 1271 GGAGCGGCGGGGAGCAAGCCAGGCGACCCAGAGCCAAAGTTTCGCCCTTCAGGACCC 1330
 DB 182 GGAGCGGCGGGGAGCAAGCCAGGCGACCCAGAGCCAAAGTTTCGCCCTTCAGGACCC 241
 QY 1331 TGAGTGCTTCGCCCACTGCTGCGATTCTACGACGCGATCTGCAATATGGAGGAGGCGAG 1390
 DB 242 TGAGTGCTTCGCCCACTGCTGCGATTCTACGACGCGATCTGCAATATGGAGGAGGCGAG 301
 QY 1391 TCCACACCCCTGCTGCACGTGGGCTGGGCCACCTTTCTGTGACGTCCCTAGGCGGTTT 1450
 DB 302 TCCACACCCCTGCTGCACGTGGGCTGGGCCACCTTTCTGTGACGTCCCTAGGCGGTTT 361
 QY 1451 TGAGGGACAGGTGCGGCAGAAAGTGGCGCATAGTGAGCCGAGAGCGCGAGCGGCCGAGCG 1510
 DB 362 TGAGGGACAGGTGCGGCAGAAAGTGGCGCATAGTGAGCCGAGAGCGCGAGCGGCCGAGCG 421
 QY 1511 CGAGGAGCCGTGGGGCGAGGAAGCCCGGGAGGCGCGGCGGCGGCGGCGGCGGAGTTC 1570
 DB 422 CGAGGAGCCGTGGGGCGAGGAAGCCCGGGAGGCGCGGCGGCGGCGGCGGCGGAGTTC 481
 QY 1571 CAAGCCAGAGGAGCGCCCGCGCCCGCCAGAGAGCCAGACCTGGGACAAAGGCGCTGGGACCGG 1630
 DB 482 CAAGCCAGAGGAGCGCCCGCGCCCGCCAGAGAGCCAGACCTGGGACAAAGGCGCTGGGACCGG 541

/lab_host="DH10B (phage-resistant)"
/note="Organ: small intestine; Vector: pCMV-SPORT6;
Site.1: NotI; Site.2: SalI; Cloned unidirectionally;
oligo-dT primed. Average insert size 1.767 kb. Library
enriched for full-length clones and constructed by Life
Technologies. Note: this is a NIH_MGC Library."
BASE COUNT 136 a 213 c 247 g 91 t
ORIGIN

Query Match 22.1%; Score 613; DB 172; Length 687;
Best Local Similarity 98.0%; Pred. No. 2.1e-133;
Matches 652; Conservative 0; Mismatches 10; Indels 3; Gaps 3;
QY 1113 CGGGAAGACGAGGAGATCTACAGGAGTTCTTTGAAGTAGCCAAATGATGTCATCCCCAAC 1232
|||||
Db 26 CGGGAAGACGAGGAGATCTACAGGAGTTCTTTGAAGTAGCCAAATGATGTCATCCCCAAC 85
|||||
QY 1173 CGGGAAGACGAGGAGATCTACAGGAGTTCTTTGAAGTAGCCAAATGATGTCATCCCCAAC 1232
|||||
Db 86 CGGGAAGACGAGGAGATCTACAGGAGTTCTTTGAAGTAGCCAAATGATGTCATCCCCAAC 145
QY 1233 CTGCTGAAGGAGGACGACGAGTCTGAGGCGGGGAGAGGCGGGGAGGAGCAAGC 1292
|||||
Db 146 CTGCTGAAGGAGGACGACGAGTCTGCTGGAGCGGGGAGAGGCGGGGAGGAGCAAGC 205
QY 1293 CAGGGCACCAGAGCAAGGTTCCGCCCTCCAGGACCTGAGTGCTCGCCACCTGCTG 1352
|||||
Db 206 CA-GGCACCAGAGCAAGGTTCCGCCCTCCAGGACCTGAGTGCTCGCCACCTGCTG 264
QY 1353 CGATTCTACGACGCGATCTGCAAAATGGGAGGAGGCGAGTCCCACGCTGTGTCACGCTG 1412
|||||
Db 265 CGATTCTACGACGCGATCTGCAAAATGGGAGGAGGCGAGTCCCACGCTGTGTCACGCTG 324
QY 1413 GGCTGGGCGACCTTTCTTGTGAGTCCCTAGGCGGTTTGTAGGAGCAGGTGGGCGAGAAG 1472
Db 325 GGCTGGGCGACCTTTCTTGTGAGTCCCTAGGCGGTTTGTAGGAGCAGGTGGGCGAGAAG 384
QY 1473 GTGCGCATAGTAGCGGAGGCGGAGGCGGCGGAGGCGGAGGCGGAGGCGGCGGAGGAA 1532
Db 385 GTGCGCATAGTAGCGGAGGCGGAGGCGGCGGAGGCGGAGGCGGAGGCGGCGGAGGAA 444
QY 1533 GCCCGGGAAGGCGGCGGCGGCGGCGGCGGCGGCGGAGGCGGAGGCGGCGGCGGCGG 1592
Db 445 GCCCGGGAAGGCGGCGGCGGCGGCGGCGGCGGCGGAGGCGGAGGCGGCGGCGGCGG 504
QY 1593 CCCAAGAGCCAGCAGTGGACAGGCGCTGGGACCGGCGGAGGCGGAGGCGGAGGCGG 1652
Db 505 CCCAAGAGCCAGCAGTGGACAGGCGCTGGGACCGGCGGAGGCGGAGGCGGAGGCGG 563
QY 1653 CCCCGAAGGCTCTGGGACTGCTGGGCGGCGGCGGCGGCGGCGGAGGCGGAGGCGGAG 1712
Db 564 CCCCGAAGGCTCTGGGACTGCTGGGCGGCGGCGGCGGCGGAGGCGGAGGCGGAGGCG 623
QY 1713 GCTCAGGTGCCAGCACCAGCAGTACCAACCGCGGAGGCGGAGGCGGAGGCGGAGGCG 1772
Db 624 GCTCAGGTGCCAGCACCAGCAGTACCAACCGCGGAGGCGGAGGCGGAGGCGGAGGCG 682
QY 1773 AGTGA 1777
|||||
Db 683 AGTGA 687

RESULT 15
BG251389
LOCUS 602364035F1 NIH_MGC_90 Homo sapiens cDNA clone IMAGE:4472432 5',
DEFINITION mRNA sequence.
ACCESSION BG251389
VERSION BG251389.1 GI:12761205
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-re@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL0293 row: e column: 09
High quality sequence stop: 681.
FEATURES
source
1..738
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4472432"
/clone_lib="NIH_MGC_90"
/tissue_type="adenocarcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/note="Organ: liver; Vector: pCMV-SPORT6; Site.1: NotI;
Site.2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.7 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
BASE COUNT 112 a 244 c 231 g 151 t
ORIGIN

Query Match 21.9%; Score 606; DB 175; Length 738;
Best Local Similarity 97.4%; Pred. No. 9.6e-132;
Matches 701; Conservative 0; Mismatches 10; Indels 9; Gaps 8;
QY 88 GCGCGCGCCACCGCGCGGCTGAAGGCGCGCCAGAGACGCTGTTCGCCG 147
Db 7 GCGCGCGCCACCGCGCGGCTGAAGGCGCGCCAGAGACGCTGTTCGCCG 66
QY 148 TGGCTCCATCGACGAGTGGTGGCTGTTCGTCGCGAGCTGGGCGGAGAGCGCG 207
Db 67 TGGCTCCATCGACGAGTGGTGGCTGTTCGTCGCGAGCTGGGCGGAGAGCGCG 126
QY 208 ACCGTGGTGTCTTCCTGGTGGCTGGGCTTCGTGGAGCATTTCTGGCTGTCAACGCG 267
Db 127 ACCGTGGTGTCTTCCTGGTGGCTGGGCTTCGTGGAGCATTTCTGGCTGTCAACGCG 186
QY 268 TCATCCTACCAAGTTCGCGAGCTCACCTTCAGCCCGCCGCGCGCGCGCGCTG 327
Db 187 TCATCCTACCAAGTTCGCGAGCTCACCTTCAGCCCGCCGCGCGCGCGCGCTG 246
QY 328 GCGGCTCACCTACTTTCGCGCGGCGACCTGTCTATCATCGCGCGCTCTATGCGCGCT 387
Db 247 GCGG-CTCACCTACTTTCGCGCGGCGACCTGTCTATCATCGCGCG-CTCTATGCGCGCT 304
QY 388 TCACGCGCGAGATCCGAGGCGCGCTGCACCTGTCCCTCTATCTCTCGAGAGGGGTGTCT 447
Db 305 TCACGCGCGAGATCCGAGGCGCGCTGCACCTGTCCCTCTATCTCTCGAGAGGGGTGTCT 364
QY 448 CCAGCGGTGAGCTGGTGAAGAAGTCTCCGATGTCATATGAAACAGCTCAGCGCGCTCT 507
Db 365 CCAGCGGTGAGCTGGTGAAGAAGTCTCCGATGTCATATGAAACAG-CTCAGCGCGCTCT 423
QY 508 ACTTCAAGAGTCCGGGCGCACATCCAGTCCCTCTTTCAGCTTCATCAGGACCAAAATGG 567
Db 424 ACTTCAAGAGTCCGGGCGCACATCCAGTCCCTCTTTCAGCTTCATCAGGACCAAAATGG 483
QY 568 ACAGTCCCGGTGTGGCGCTTTGCTGTGGTGGGCGCTGCCAGGCGCTGGGTCTCGCGATG 627
Db 484 ACAGTCCCGGTGTGGCGCTTTGCTGTGGTGGGCGCTG-CAGGCGCTGGGTCTCGCGATG 542

Qy 628 TCCACCTCGCCCTGCTGAGGATCATGCTGGGTAGTGTGTTGGCCCAATGGGAGCAGA 687
Db 543 TCCA-CTCGCCCTGCTGAGGATCATGCTGGGTAGTGTGTTGGCCCAATGGGAGCAGA 601
Qy 688 CAGCTGAGGTACCTGCGACGGCAAGGGCAACGAGGACCGAGGGCCAGACAGTCAATG 747
Db 602 CAGCTGAGGTACCTGCGACGGGACGGCAACGAGGACCGAGGGG-CAGACAGTC-ATG 659
Qy 748 CCGGTGTGGCTGAGCGGAGCTGGCTGTACCTGAAAGGATCATACATGCGCTGTGACCGCA 807
Db 660 CCGGTGTGGCTGAGCGGAGCTGGGTGTTCCTGAAGGG--TCATCATGCGCTGTGACCGCA 717

Search completed: October 14, 2001, 03:45:17
Job time: 28023 sec

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	2772	100.0	2772	5	US-09-948-941-5	Sequence 5, Appli
2	2772	100.0	2772	5	US-09-980-337-1	Sequence 1, Appli
3	2707.8	97.7	2767	5	US-09-948-941-135	Sequence 125, App
4	2580.2	93.1	3534	8	US-60-278-561-1743	Sequence 1743, Ap
5	2046.6	73.8	2168	6	US-09-758-466-258	Sequence 258, App
6	1312.2	47.3	9180	5	US-09-380-337-3	Sequence 3, Appli
7	1312.2	47.3	9180	7	US-09-880-107-3461	Sequence 3461, Ap
8	1253.2	45.2	8185	6	US-09-760-475-3945	Sequence 3945, Ap
9	1253.2	45.2	11190	5	US-09-948-941-529	Sequence 529, App
10	1253.2	45.2	11205	5	US-09-948-941-649	Sequence 649, App
11	529.4	19.1	628	8	US-60-253-651-22376	Sequence 22376, A
12	488.4	17.6	639	7	US-09-864-761-22111	Sequence 22111, A
13	473.4	17.1	519	6	US-09-878-178-1236	Sequence 1236, Ap
14	452	16.3	464	7	US-09-864-761-5337	Sequence 5337, Ap
15	447	16.1	3534	8	US-60-278-561-1743	Sequence 1743, Ap
16	446	16.1	2121	1	PCN-US01-14827-5297	Sequence 5297, Ap
17	446	16.1	2770	1	PCN-US01-14827-388	Sequence 388, App
18	440	15.3	479	6	US-09-898-888-29321	Sequence 29321, A
19	415.2	15.0	518	8	US-60-253-652-892	Sequence 892, App
20	397.6	14.3	599	8	US-60-253-651-22088	Sequence 22088, A
21	376.2	13.6	433	8	US-60-252-833-33948	Sequence 33948, A
22	368	13.3	429	8	US-60-252-833-33900	Sequence 33900, A
23	366	13.2	414	8	US-60-252-833-33919	Sequence 33919, A
24	323	11.7	447	6	US-09-909-629-32588	Sequence 32588, A
25	318	11.5	391	6	US-09-904-703-12130	Sequence 12130, A

QY	361	CTATCATCGCGCCCTCTATGCCCGCTTACCGCCAGATCCGAGGCGCGGTGCACTGT	1441	TAGCCCGTTTTGGAGGACAGGTGCGGAGAGGTGCCATAGTAGCCGAGAGCCGAGG	1500
Db	361	ctatcatcgcgccctctaTgcgcgttcaaccgccaagtcgagcgcgctcgacctgt	1441	taggcccgttttgaggagcaggTgcgagaggtTgcgcataTgTgagccgagagccgag	1500
QY	421	CCCTCTATCTCAGAGAGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAGGTCTCCGATG	1501	CGCCGAGGCGCGAGGAGCCGTGGGGCAGGAAGCCGGGAAGCCGCGGGGGCCAC	1560
Db	421	ccctctatctcagaaagggTgtctccagccgtTgagctTgTgaaagaggtctccgatg	1501	cggccagggccgagggagccgtTggggcgaggaagcccggaagcgcgcgggggccac	1560
QY	481	TCATATGAACAGCTCAGCGCTCTACTTCAAGGATCGGGCCACATCCAGTCCCTCT	1561	GGCGGGAGTCCAAGCCAGAGAGCCCGCCGCCCAGAAAGCCAGCACTGCACAAAGGCC	1620
Db	481	tcatatggaacagcctcagccgtccctcacttcaaggatcgggcccatccagtcctct	1561	ggcgggagttccaagccaagagagcccccgcgcccgaagaagccaactTgacaagggcc	1620
QY	541	TCAGCTTCATCAGAGCACCANAATGGACAGCTCCGGTGTGGCTTTGCTGTGGTGGG	1621	TGGGCACCCGCCAGGGTGCAGTGTCAAGACCCGCCCGGAAAGCCCTTGGGACTGTCTG	1680
Db	541	tcagcttcatacaagcagcaaaaattgacagctccggTgtggcctttgtctgtgtggg	1621	TgggcaaccggccagggTgcagTgtTcaggacccccgggaagccctctTggactgtcgctg	1680
QY	601	CTGCGCAGGCGCTGGGTCTCGGGATGTCCACTCGCCCTGTCTGAGGATCATGCTCTGG	1681	GCACAGCCCGCAGGCCCTGAAGTGGCAGCAGCGTCAAGTGGCCAGCACCCGACATCAC	1740
Db	601	cctgccagggccctgggtctccggatgtccacctgcacctgtctgaggaTcatgtctggg	1681	gcaagcccagggccctTgaaggTggcagcacggctcaggtTgccaagcaccgcgagcaTcac	1740
QY	661	TAGTGTGGGCCCCAATGGGAGCAGACAGCTGAGGTCACTGGCAGCGCAAGGGCAACG	1741	CACCGCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGAATGAAGGCGATGAAGGAGC	1800
Db	661	tagTgtttggggcccaatggggagcagacagctgaggtTcacctTggcagcgcaagggcaacg	1741	caccgcggagggTccaagtgtctcaactttccagagtgaagaTgaagggcatgaagagagc	1800
QY	721	AGGACCGAGGGCCAGACAGTCAATCCCGTGTGGCTGAGCGGAGCTGGCTGTACCTGA	1801	TGCTGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCCAACTCAGCGCACAGTCGC	1860
Db	721	aggaccgagggccagacagtaaatcccggtTgtggctgagcgagctgctgtacctga	1801	TgtgtgtggccaccaagTcaactcgagcgccaTcaagctTgaaactcacgTcacgTcgc	1860
QY	781	AAGGATATACATCGCTGTGACCGCAAGATGAGGTGGCGTTCATGGTGTGGCCATCA	1861	AAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTTCTCTCA	1920
Db	781	aaggatatacatgctgtTgaccgcaagatgaggtTggcgtTcaatggtTgtgtgccaTca	1861	agTgcagaTgaagaagcagaaggtTccaccctctagTgactacactctgtTcttctctca	1920
QY	841	ACCCTTCATTAGCTGCACACCACTCGCTGGAGCTTCTGCACTGCAGCAGCAAGCTGC	1921	AGGGCAGCCCAAGGCTCTGAACTACTGGGAGCTTCGGACCCCTTGTGGGGACCCAGG	1980
Db	841	acccttcattgacctgcaacagagctcgctggagctTctgcaagctgcagcagaagctgc	1921	agcgcaagcacaagggcctctgaaactactgggactTcggaacgctTgtgggagcccgag	1980
QY	901	TCTGGCTGCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCCTTAGGGAACCTGG	1981	CTCGCCTTAGTCCCGCACTCTGAGCGCCATGTCTGCCCGCCAGCCCAAGGGGACAGGC	2040
Db	901	tctTgctTgctctatgactggagacatctgaaaggTaccccaatggccttaggaaacctg	1981	ctcgcccttagtcccccaactctgagcccatgttctgccccccagccccaaaggggagcggc	2040
QY	961	CAGATCTAGGAGCTGGAGCCACCCCTTGCGCGCCAGACCCCTACCCCTCTACCACA	2041	CTCACTCTACCCAAACCCCTAGGTTCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGA	2100
Db	961	cagatctagaggagctggagccaccctctggccggcagaccactcacacctctaccaca	2041	ctcaactctaccacaaccccttaggtTccccgtTccccgagTacaagctgtatcaaacccacga	2100
QY	1021	AGGCACTTGCTCAGCCAAGACCTACTATCGGATGAACATCTACCCCTTACATGTACC	2101	TTTTTCTCCAGCTCAGAACCAGGCTCTGCCCGCCAGTCTGTAGATAATAGTCTCTCTCC	2160
Db	1021	aggcattgctcagcaagacctactatcggaTgaacacatacctaccctatactgtacc	2101	TttctccagctcagaacccagggctctgTgcccagTcgTtagaataTaggTctctctcc	2160
QY	1081	TGGTGTGCTACCACTGTGCAACCGCAATGTGCGGGAAGCCCTGCAGGCTTGGCGGACA	2161	CAGNAATCCCGAGCGGCCAATGGAAACCTCAGCTGGGTCTCTAATTACCACTCTTTAAGG	2220
Db	1081	TggTggtTaccactTgcgaaccgcaatTgcgggaagccctgcagggcctTgpgcgga	2161	cagaatccccagcgcgcaatTggaacactcacgtTgggtTcctaatTaccagTctttaaagg	2220
QY	1141	CGGCCACTGTATCAGGACTACAACTACTTCCCGGGAAGACGAGGATCTACAAAGAGT	2221	CCACGCCCTTAGAAACCAAGCTCCTCTCGGAACCGCTCACCTTAGAGCCAGACCAACGT	2280
Db	1141	cggccaactgtcaatcaggaactacaactactTgcgggaagcgagagagatctacaaggat	2221	cccaagccccTagaaaccccaagctctctctcTggaaccctcacctcagagccagacaacgt	2280
QY	1201	TCTTTGAAGTAGCAATGATGTCTATCCCAACCTTCGCTGAAGGAGGCAGAGCTTGTCTGG	2281	TACTCAGGGCTCTCTCCAGCTTGTAGGAGCTGAGGTTTTCACCCCTTAACCCCAAGGAGCAC	2340
Db	1201	tctTtgaagtagccaaTgaTgtcaTccccaaactTgtTgaaggaggaagcaagctTgctgg	2281	TactcagggctctctcccgactTgttaggagctgaggtTtccaccttaaccaagggagcac	2340
QY	1261	AGCGGCGAGGAGCGCGGGGAGCAAGCCAGGACCCAGGACCAAGGTTTCGCCCC	2341	AGGTCCACCTTCCAGCCCGGGGAGCCCTAGGACCACTCAGGCCCTTAGGAGTATATTTCGGC	2400
Db	1261	aggcgcgagagagcgcgggggagcaaaagcagggTcaaccagagccaaggtTcccgccc	2341	aggTccccacctccagcccgggagccttaggaccactcagcccttaggataTattctcgc	2400
QY	1321	TCCAGGACCTGAGTGTGCTGCCACCTGTGCGATTTCTACGAGCGGATCTGCAAAATGGG	2401	ACTTTCAGAAATTCATATCTTTCGGAATCCAAGTCCCTCGCCCAATAACTTTCAGTCTCTGC	2460
Db	1321	tccaggacctgagTgtTgcggccacctgtgcgatTctaagacgggcatctgcaaaTggg	2401	acttTcagaatttccataTctTgcgaatccaagctccccTgcggcccaataaactTcagTctctgc	2460
QY	1381	AGGAGGCACTCCCAAGCGCTGTGCTGCACTGGCTGGGCGACACTTCTTGTGCACTGCC	2461	TTCAGAAATTTGGAATTCCTAGTTTCTCTCTCTTCTGATCCCGAGTCTGGGACACAAAC	2520
Db	1381	aggaggcagTccccagccctgtgctgcagTgggctgggcccaccttctTgtgagTccc	2461	TtccagaatttggaatTcctagTttctctctctctctTatccccgagTctTgggacacaaac	2520
			2521	TCCGCCCCCGAGCCTATGAGCATCTCTGAGCCCCCGCCCTTCTCTGACGNAACCTGGCCCCGG	2580

Db 2521 tccgccccagcctatgagcatcctgagcccccctctctcctgagcaactgcccccg 2580
QY 2581 ATCAGAGCAGGACCTCCCTTCCGACCTCTGGGAACCTCCAGAGGTCCAGGCCATCTCG 2640
Db 2581 atcagagcagcctccctccgacctctggaacctccagagcctccagcccactcg 2640
QY 2641 GAGCATCCCGAGGAATTCGACAGGGGTAGGAGTGGGTGACAGAGCCTGATCTCTT 2700
Db 2641 gagcatcccgaggaaatctcgagagggttagagtggtgagagcctgctctt 2700
QY 2701 CCGTCTTGTACATAGATTATTTTCAGTTCCTCAAGAAAGATGAATTTGTTAAAA 2760
Db 2701 cctgtttgtacatagattatttttcagttccaagaagatgaatacatttgttaaa 2760
QY 2761 AAAAAAAAAA 2772
Db 2761 aaaaaaaaaa 2772

RESULT 2

US-09-380-337-1
: Sequence 1, Application US/09380337
: GENERAL INFORMATION:
: APPLICANT: Chandrasekharappa, Settara C.
: Guru, Siradanahalli C.
: Manickam, Pachaiappan
: Collins, Francis S.
: Emmert-Buck, Michael R.
: Debelenko, Larisa V.
: Lubensky, Irina A.
: Liotta, Lance A.
: Agarwal, Sunita K.
: Spiegel, Allen M.
: TITLE OF INVENTION: MEN1, the Gene Associated with Multiple
: Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses
: Thereof

NUMBER OF SEQUENCES: 28
CORRESPONDENCE ADDRESS:
: ADDRESS: Townsend and Townsend and Crew LLP
: STREET: Two Embarcadero Center, Eighth Floor
: CITY: San Francisco
: STATE: California
: COUNTRY: USA
: ZIP: 94111-3834
COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/380,337
: FILING DATE: 09-Mar-2000
: CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 60/040,269
: FILING DATE: 05-MAR-1997
: APPLICATION NUMBER: WO PCT/US98/04258
: FILING DATE: 04-MAR-1998
ATTORNEY/AGENT INFORMATION:
: NAME: Lockyer, Jean M.
: REGISTRATION NUMBER: 44,879
: REFERENCE/DOCKET NUMBER: 015280-315100US
TELECOMMUNICATION INFORMATION:
: TELEPHONE: (415) 576-0200
: TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2772 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA

FEATURE:
NAME/KEY: CDS
LOCATION: 111..1940
OTHER INFORMATION: /product= "human menin"
FEATURE:
NAME/KEY: exon
LOCATION: 1..87
FEATURE:
NAME/KEY: exon
LOCATION: 88..555
FEATURE:
NAME/KEY: exon
LOCATION: 556..764
FEATURE:
NAME/KEY: exon
LOCATION: 765..893
FEATURE:
NAME/KEY: exon
LOCATION: 894..934
FEATURE:
NAME/KEY: exon
LOCATION: 935..1022
FEATURE:
NAME/KEY: exon
LOCATION: 1023..1159
FEATURE:
NAME/KEY: exon
LOCATION: 1160..1295
FEATURE:
NAME/KEY: exon
LOCATION: 1296..1460
FEATURE:
NAME/KEY: exon
LOCATION: 1461..2764
SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-380-337-1
Query Match 100.0%; Score 2772; DB 5; Length 2772;
Best Local Similarity 100.0%; Pred No. 0;
Matches 2772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTGTCCGGAGCCGCGACCTAGAGATCCAGAAAGCCACAGCGAGCGCGCCGCCGCTGA 120
Db 1 GGTGTCCGGAGCCGCGACCTAGAGATCCAGAAAGCCACAGCGAGCGCGCCGCCGCTGA 120
QY 61 ACTATTTCAGAGCTCTCCGGGCGAGGGGCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCG 120
Db 61 ACTATTTCAGAGCTCTCCGGGCGAGGGGCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCG 120
QY 121 AGCGCGCCCAAGAGAGCGCTGTCCCGCTGCGCTCCATCGACGACGTGGTGGCTGTTG 180
Db 121 AGCGCGCCCAAGAGAGCGCTGTCCCGCTGCGCTCCATCGACGAGCGTGGTGGCTGTTG 180
QY 181 CTCCCGAGCTGGGCGCGAGAGCGGACCTGGTGTCTCTTCTTCTTCTTCTTCTTCTTCTT 240
Db 181 CTCCCGAGCTGGGCGCGAGAGCGGACCTGGTGTCTCTTCTTCTTCTTCTTCTTCTTCTT 240
QY 241 TGGAGCATTTTCTGGGTGTCAACCGGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCC 300
Db 241 TGGAGCATTTTCTGGGTGTCAACCGGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCC 300
QY 301 AGCGCGCCCGCCCGCCCGCGCGCTGGCGGCTCAGCTACTTCTTCTTCTTCTTCTTCTTCTT 360
Db 301 AGCGCGCCCGCCCGCCCGCGCGCTGGCGGCTCAGCTACTTCTTCTTCTTCTTCTTCTTCTT 360
QY 361 CTATCATCGCGCCCTCTATGCGCGTTTACCGGCCAGATCCGAGGCGCGCTGACCTGT 420
Db 361 CTATCATCGCGCCCTCTATGCGCGTTTACCGGCCAGATCCGAGGCGCGCTGACCTGT 420
QY 421 CCCTCTATCTCGAGAGGGGTGTCTCCAGCGGTGAGCTGCTCAAGAGGTCTCCGATG 480
Db 421 CCCTCTATCTCGAGAGGGGTGTCTCCAGCGGTGAGCTGCTCAAGAGGTCTCCGATG 480

Qy	481	TCATATGGAAACAGCCTTCAGCCGCTCCTACTTCTCAAGGATCGGGGCCACATCCAGTCCCTCT	540
Db	481	TCATATGGAAACAGCCTTCAGCCGCTCCTACTTCTCAAGGATCGGGGCCACATCCAGTCCCTCT	540
Qy	541	TCAGCTTCATACAGGCACCAAAATTTGGACAGCTCCGGTGTGGCCCTTTGCTGTGGTTGGGG	600
Db	541	TCAGCTTCATACAGGCACCAAAATTTGGACAGCTCCGGTGTGGCCCTTTGCTGTGGTTGGGG	600
Qy	601	CTTGCCAGGCCCTGGGTCTTCGGGATGTCACACTCGCCCTGCTGAGGATCATGCTCGTGG	660
Db	601	CTTGCCAGGCCCTGGGTCTTCGGGATGTCACACTCGCCCTGCTGAGGATCATGCTCGTGG	660
Qy	661	TAGTGTTTGGGCCCAATGGGGAGCAGACAGCTGAGGTCACTTGGCAGCGCAAGGCAACG	720
Db	661	TAGTGTTTGGGCCCAATGGGGAGCAGACAGCTGAGGTCACTTGGCAGCGCAAGGCAACG	720
Qy	721	AGGACCGCAGGGGCCAGACGATCAATGCGGTGTGGCTGTGACGGAGCTGGCTGTACCTGA	780
Db	721	AGGACCGCAGGGGCCAGACGATCAATGCGGTGTGGCTGTGACGGAGCTGGCTGTACCTGA	780
Qy	781	AAGGATCATACATGCGCTGTGACCCCAAGATGGAGTGGCGTTTCATGCTGTGCCATCA	840
Db	781	AAGGATCATACATGCGCTGTGACCCCAAGATGGAGTGGCGTTTCATGCTGTGCCATCA	840
Qy	841	ACCTTTCCATTGACCTGTCACACCGACTCGCTGGAGCTTCTTGCAGCTGCAGCAAGCTGC	900
Db	841	ACCTTTCCATTGACCTGTCACACCGACTCGCTGGAGCTTCTTGCAGCTGCAGCAAGCTGC	900
Qy	901	TCTGGCTGCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCCCTTAGGGAACCTGG	960
Db	901	TCTGGCTGCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCCCTTAGGGAACCTGG	960
Qy	961	CAGATCTAGAGGAGCTGGAGCCCCACCCCTGGCCGGCCAGACCCACTCACCCCTTACCACA	1020
Db	961	CAGATCTAGAGGAGCTGGAGCCCCACCCCTGGCCGGCCAGACCCACTCACCCCTTACCACA	1020
Qy	1021	AGGCGATTGGCTTCAGCACAAGACTTACTATCGGGATGAACATCTACCCCTACATGTACC	1080
Db	1021	AGGCGATTGGCTTCAGCACAAGACTTACTATCGGGATGAACATCTACCCCTACATGTACC	1080
Qy	1081	TGGCTGGCTACCACTGTCGCAACCGCAATGTGCGGGAAGCGAGGAGATCTACAAGGAGT	1140
Db	1081	TGGCTGGCTACCACTGTCGCAACCGCAATGTGCGGGAAGCGAGGAGATCTACAAGGAGT	1140
Qy	1141	CGGCCACTGTCATCCAGGACTACAACACTACTCGCGGGAAGACGAGGAGATCTACAAGGAGT	1200
Db	1141	CGGCCACTGTCATCCAGGACTACAACACTACTCGCGGGAAGACGAGGAGATCTACAAGGAGT	1200
Qy	1201	TCPTTTGAAGTAGGCAATGATGTATCCCCAACCTGCTCAAGGAGGCAGCAGCTTGTCTGG	1260
Db	1201	TCPTTTGAAGTAGGCAATGATGTATCCCCAACCTGCTCAAGGAGGCAGCAGCTTGTCTGG	1260
Qy	1261	AGCGGGCGAGGAGCGCGGGGGAGCAAGCAGGGGCACCCAGAGCCAGGTTTCGCGCC	1320
Db	1261	AGCGGGCGAGGAGCGCGGGGGAGCAAGCAGGGGCACCCAGAGCCAGGTTTCGCGCC	1320
Qy	1321	TCAGAGACCTGAGTGTCTGCGCCACCTGCTCGGATTCCTACAGCGGCATCTGCAAAATGGG	1380
Db	1321	TCAGAGACCTGAGTGTCTGCGCCACCTGCTCGGATTCCTACAGCGGCATCTGCAAAATGGG	1380
Qy	1381	AGGAGGCGACTCCACCGCTGTCTGCAGTGGCTGGGCCACCTTTCTTCTGCAGTCCC	1440
Db	1381	AGGAGGCGACTCCACCGCTGTCTGCAGTGGCTGGGCCACCTTTCTTCTGCAGTCCC	1440
Qy	1441	TAGGCCCTTTTGGAGGACAGGTGCGGCAGAAAGGTGCGCATAGTGAGCCGAGAGCCGAGG	1500
Db	1441	TAGGCCCTTTTGGAGGACAGGTGCGGCAGAAAGGTGCGCATAGTGAGCCGAGAGCCGAGG	1500
Qy	1501	CGGCCGAGGCCGAGGAGCGTGGGGCGAGGAAGCCCGGGGAAGCGCGGGGGGGCCAC	1560
Db	1501	CGGCCGAGGCCGAGGAGCGTGGGGCGAGGAAGCCCGGGGAAGCGCGGGGGGGCCAC	1560

Qy	1561	GGGGGAGTCCAAAGCCAGAGAGAGCCCGCCCGCCCAAGAGCCAGACTGGACAGAGGGCC	1620
Db	1561	GGGGGAGTCCAAAGCCAGAGAGAGCCCGCCCGCCCAAGAGCCAGACTGGACAGAGGGCC	1620
Qy	1621	TGGGCACCGGCCAGGGTGCAGTGTCTAGGACACCCCGGAGCCCTCTGGGACTGTCTGCTG	1680
Db	1621	TGGGCACCGGCCAGGGTGCAGTGTCTAGGACACCCCGGAGCCCTCTGGGACTGTCTGCTG	1680
Qy	1681	GCACAGCCCGAGGCCCTGAAGTGGCAGCACGGCTTCAGGTGCCAGCACCCCGCAGCATCAC	1740
Db	1681	GCACAGCCCGAGGCCCTGAAGTGGCAGCACGGCTTCAGGTGCCAGCACCCCGCAGCATCAC	1740
Qy	1741	CACGGCGGAGGTCCAGTGTCTACTTTTCCAGAGTGAGAGATGAAGGCGATGAAGGAGC	1800
Db	1741	CACGGCGGAGGTCCAGTGTCTACTTTTCCAGAGTGAGAGATGAAGGCGATGAAGGAGC	1800
Qy	1801	TGCTGGTGGCCACCAAGATCAACTCGAGGGCCATCAAGCTGCAACTCAGGCGACAGTCCG	1860
Db	1801	TGCTGGTGGCCACCAAGATCAACTCGAGGGCCATCAAGCTGCAACTCAGGCGACAGTCCG	1860
Qy	1861	AAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTACACTCTGTCTTTCTCTCA	1920
Db	1861	AAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTACACTCTGTCTTTCTCTCA	1920
Qy	1921	AGCGGCAGCCAAAGGCCCTCTGAACTACTTGGGGACTTCGGACCCCTGTGTGGGGACCCAGG	1980
Db	1921	AGCGGCAGCCAAAGGCCCTCTGAACTACTTGGGGACTTCGGACCCCTGTGTGGGGACCCAGG	1980
Qy	1981	CTCGGCCTTAGTCCCCCACTCTGAGCCCATGTTCTGCCCCAGCCCNAAAGGGGACAGGC	2040
Db	1981	CTCGGCCTTAGTCCCCCACTCTGAGCCCATGTTCTGCCCCAGCCCNAAAGGGGACAGGC	2040
Qy	2041	CTCAGCTCTACCCAAACCCCTAGSTTCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGA	2100
Db	2041	CTCAGCTCTACCCAAACCCCTAGSTTCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGA	2100
Qy	2101	TTTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCCAGTCTGTAGAATATAGGTCTCTTCTCC	2160
Db	2101	TTTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCCAGTCTGTAGAATATAGGTCTCTTCTCC	2160
Qy	2161	CAGAACTCCAGCCGGCCAAATGGAAACCTCAGCTCGGTGCTCTAATTTACGAGTCTTTAAAGG	2220
Db	2161	CAGAACTCCAGCCGGCCAAATGGAAACCTCAGCTCGGTGCTCTAATTTACGAGTCTTTAAAGG	2220
Qy	2221	CCAGAGCCCTAGAAACCCCAAGCTCTCTCTCGGAACCGCTCAGCTTAGAGCCAGACCAACGT	2280
Db	2221	CCAGAGCCCTAGAAACCCCAAGCTCTCTCTCGGAACCGCTCAGCTTAGAGCCAGACCAACGT	2280
Qy	2281	TACTCAGGGCTCTCCAGCTTGTAGGAGCTGAGGTTTTCAACCTTAACCCAAAGGAGCAC	2340
Db	2281	TACTCAGGGCTCTCCAGCTTGTAGGAGCTGAGGTTTTCAACCTTAACCCAAAGGAGCAC	2340
Qy	2341	AGGTCCCACTCCAGCCGGGGAGCCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGC	2400
Db	2341	AGGTCCCACTCCAGCCGGGGAGCCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGC	2400
Qy	2401	ACTTCAGAAATTCATATCTTGGAAATCCAAAGCTCCCTGCCCAATAACTTTCAGTCTCTGC	2460
Db	2401	ACTTCAGAAATTCATATCTTGGAAATCCAAAGCTCCCTGCCCAATAACTTTCAGTCTCTGC	2460
Qy	2461	TTCCAGAAATTTGGAATTCCTAGTCTTCTCTCTCGTATCCCGAGTCTGGGACACAAAC	2520
Db	2461	TTCCAGAAATTTGGAATTCCTAGTCTTCTCTCTCGTATCCCGAGTCTGGGACACAAAC	2520
Qy	2521	TCCGCCCCAGCCTATGAGCATCCTGAGCCCGCCCTCTTCTGTACAGAACTGCGCCCGG	2580
Db	2521	TCCGCCCCAGCCTATGAGCATCCTGAGCCCGCCCTCTTCTGTACAGAACTGCGCCCGG	2580
Qy	2581	ATCAGACGAGACCTCCCTCCGACCCCTCTGGGAACCTCCCGAGGTCCACGCCATCTCG	2640
Db	2581	ATCAGACGAGACCTCCCTCCGACCCCTCTGGGAACCTCCCGAGGTCCACGCCATCTCG	2640
Qy	2641	GAGCATCCCGGAGGAAATCTGCAGAGGGGTATGAGTGGGTGACAAAGAGCCTGTATCTCTT	2700


```
|||||
Db 2641 GAGCATCCGGAGGAAATCTGCAGAGGGTTAGGAGTGGGTGACAAAGAGCCTGATCTCT 2700
QY 2701 CCGTGTTCACATACATTTATTTTTCAGTCCAAAGAAAGATCAATACATTTTGTAA 2760
Db 2701 CCGTGTTCACATACATTTATTTTTCAGTCCAAAGAAAGATCAATACATTTTGTAA 2760
QY 2761 AAAAAAAAAA 2772
Db 2761 AAAAAAAAAA 2772
```

```
RESULT 3
US-09-948-941-125
; Sequence 125, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948,941
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 125
; LENGTH: 2767
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-125
```

```
Query Match 97.7%; Score 2707.8; DB 5; Length 2767;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 2764; Conservative 0; Mismatches 2; Indels 5; Gaps 5;

QY 1 GGTGTCGGAGCGCGACCTAGAGATCCAGAGGACACAGCGAGCGCGCGCGCGCC 60
Db 1 ggtgtcggagcgcgacctagatccagagggacacagcagcagcagcgcgcgcgc 60
QY 61 ACTATTTCCAGGCTCTGCGGGGAGGGGCGCGCGCGCGCGCGCGCGCGCGCGCG 120
Db 61 actatttcaggctctgcggggagggcgcgcgcgcgcgcgcgcgcgcgcgcg 120
QY 121 AGCGCGCCAGAGAGAGCTCTCCGCTGCGCTCCATCGAGAGAGTGTGCGCTG 180
Db 121 agcgcgccagagagagcgtctcccgctgcgctccatcgagcagcgtggtgc 180
QY 181 CTGCGAGCTGGGCGGAGAGAGCGGACCTGGTGTCTCTTCTTGGTGTGCGGCT 240
Db 181 ctgcgagctggcgagagagcggaacctggtgctcttcttctggtggtcg 240
QY 241 TGAGAGATTTCTGGCTGTCAACGGGTATCCCTACCAAGCTTCCGAGCTCACCT 300
Db 241 tggagatcttctggtgtcaacgggtatccctaccaagcttccgagctcacct 300
QY 301 AGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 360
Db 301 agcgcgcccgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcgcg 360
QY 361 CTATCATCTGCGCGCGCTATCTCCCGCTTTCACCGCCAGATCCGAGGCGCG 420
Db 361 ctatcatcgcgcgcctctatgcgcgcttccacgcgcgcgcgcgcgcgcgc 420
QY 421 CCCTCTATCTCGAAGAGGGGTGTCTCCAGCGGTGAGCTGGTGAAGAGGTCTCC 480
Db 421 ccctctatctcgagaaggggtgtctccagcgtgagctggtgaagaaggtctcc 480
QY 481 TCATATGGAACAGCTCAGCGCTCTACTTCAAGGATCGGGCGCGCGCGCGCGCT 540
Db 481 tcatatggaacagcctcagcgctctacttcaaggatcgggcgccatccgctct 540
```

```
QY 541 TCAGCTTCATACAGGCACCAAAATGGACAGCTCCGTTGGCTTGTGCTGTTGGG 600
Db 541 tcagcttcacacaggcacaaaattggacagctccggtgtgaccttctgtgtgtgg 600
QY 601 CTTGCCAGGCCCTGGGTCTCCGGGATGTCACCTCGCCCTGCTGAGGATCATGCT 660
Db 601 cttgccagccctgggtctccgggatgtccacctcgccctgtctgaggtcatgct 660
QY 661 TAGTGTGGGCCCAATGGGGAGCAGACAGCTGAGGTCACTGGCAGCGCAAGGCA 720
Db 661 tagtgtttggggcccaatggggagcagacagctgaggtccacctggcagcgca 720
QY 721 AGGACCGCAGGGGCCAGACAGTCAATGCCGGTGTGGCTGAGCGAGCTGCTGCT 780
Db 721 aggaaccgcaggggccaagacagtcaatgcccgtgtggtgagcgagctggtct 780
QY 781 AAGGATCATACATGCGCTGTGACCGCAAGATGGAGTGGGCTTCATGCTGTGCC 840
Db 781 aaggatcatacatgctgtgacccgaagatggaggtggttcattggtgtggtcat 840
QY 841 ACCCTTCCATTGACCTGCACACCGAGCTCGCTGGAGCTTCTGCAGCTGCAGCA 900
Db 841 acccttcattgacctgcacacagcactcgtgagcttctgagctgcagcagaag 900
QY 901 TCTGGCTGCTCTATGACCTGGGACATCTGAAAGAGTACCCCATGGCTTAGGAA 960
Db 901 tctggctgctctatgacctgggacatctgaaaggtaccacctggccttaggaa 960
QY 961 CAGATCTAGAGGAGCTGGAGCCCAACCTCGCGCGCGCGAGACCCTACCTCTAC 1020
Db 961 cagatctagaggagctggagccaccctggcgcgcgagaccacctacccttacc 1020
QY 1021 AGGCGATTGCTCAGCAAGACCTACTATCGGGATGAACACATCTACCCCTACAT 1080
Db 1021 agggcattgctcagcgaagacactactatcggatgaacacatctacccttacc 1080
QY 1081 TGCTGGCTACCACTGTGCAACCGCAATGTGCGGGAGCCCTGCAGGCTTGGGCG 1140
Db 1081 tggctggctaccactgtcgcaacgcgaatgtcgggagccctgcagcctggcg 1140
QY 1141 CGGCCACTGTATCCAGGAGTCAACTACTTCCCGGGAGAGGAGAGATCTACAG 1200
Db 1141 cgggcactgtcatccaggactcaactactgcgggagcagcagagagatctac 1200
QY 1201 TCTTTGAAGTAGCAATGATGTCTCATCCCAACCTGCTGAAGGAGGCGAGCT 1260
Db 1201 tctttgaagttagcaatgatgtcatcccaacctgtcgaaggagcagcagct 1260
QY 1261 AGCGGGCGAGGAGCGCGCGGGAGCAAAAGCCAGGCGACCCAGAGCCAGGTT 1320
Db 1261 agcgggcgagagcgcgcggggagcaaaagccagggcaccagagcgaaggtt 1320
QY 1321 TCCAGGACCTGTAGTGTCTCGCCACCTGCTGCGATTCAGGAGGCTCTGCAAT 1380
Db 1321 tcaggacctgtgctctcgccacctgtcgatctctacgagcagcagcagctgt 1380
QY 1381 AGGAGGCGAGTCCACACGCTGTCTGACGCTGGGCTGGGCGACCTTCTTCT 1440
Db 1381 aggaaggcagtcaccaagcgtgctgcaagtgggctggcgccaccttcttct 1440
QY 1441 TAGCCCTTTTGGAGGACAGTGTGCGGAGAGTGTGCGCATAGTAGCGGAGCC 1500
Db 1441 taggccttttgaggacaggtgcggcagaaggtgcgcatagtgagcgcagag 1500
QY 1501 CGGCCGAGGCGGAGGAGCGGTGGGCGGAGGAGAGCCCGGGGAGGCGCGGGCC 1560
Db 1501 cgggcgagggcgagagcgtggtggcgaggaagcccgggagggcgcgcgggcc 1560
QY 1561 GCGGGGAGTCCAAAGCAGAGGAGCGCCCGCGCCCAAGAGCCAGCACTGTGCA 1620
Db 1561 ggcggggagtcgaagcagagagagagcccccgccgccccgaagcagcagcag 1620
QY 1621 TGGCACCGCGCGAGGGTGCAGTGTGAGGAGCCCGCGGAGAGCCCTCTCTG 1680
```

Db 1621 tgggacccgacagggtgacagtgcagacccccccggaagcctcctgggactgcgtg 1680
Qy 1681 GCACAGCCGAGGCGCTCAAGGTGGCAGCAGCGTCCAGCAGCCCGCAGCATCAC 1740
Db 1681 gcacagcccgaggccctgaaggtggcagcagcgctcaggtgcccagcccgacacac 1740
Qy 1741 CACGCGGAGGGTCCAGTGTCTACTTCCAGAGTGAAGATGAAGGCGATGAAGGAGC 1800
Db 1741 caccgcggaggggtccagtgctcacttccagagtgagagagtgaaagggcagagc 1800
Qy 1801 TGCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCG 1860
Db 1801 tgcctgggtggccaccagatcaactcagcgcacatcaagctgcaactcacggcacagtcgc 1860
Qy 1861 AAGTGCAGATGAAGAACAGAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTCTCTCA 1920
Db 1861 aagtgcagatgaagaagcagaaagtccaccctctgactacactctgtcttctctca 1920
Qy 1921 AGCGCAGCGCAAGGCGTCTGAAGTACTGGGGAGCTTCGGACCGCTTGTGGGGACCCAGG 1980
Db 1921 agcgagcagcaaaagccctctgaactactcgggacttcggacccgtgtggggacccag 1980
Qy 1981 CTCGG-CCTTAGTCCCCAACTGTAGGCCATGTCTGCCCGCCAGCCCAAGGGGACAGG 2039
Db 1981 ctccgccccttagtcccccaactctgagcccatgtctcgccccagccccaaaggagcag 2040
Qy 2040 CCTCACCTCTACCCAAACCCCTAGCTTCCGGTCCGGAGTACAGTCTGTATCAAAACCCAGC 2099
Db 2041 cctcacctctacccaaacccctaggtctcccggtcccgagtagcagctgtatcaaacccag 2100
Qy 2100 ATTTTCTCAGCTCAGAACCCAGGCGTCTGCCCGAGCTGTAGAAATATAGTCTCTTCTC 2159
Db 2101 atttctccagctcagaacccagggctctgccccagctgttagaatatagctctcttc 2160
Qy 2160 CCAGAAATCCAGCGCGCAATGGAACCTCAGCGTGGGTCTTAATACAGCTCTTTAAAG 2219
Db 2161 ccagaaatccagcggcgccaatggaaacctcagctgggtcttaataatccagctctttaa 2220
Qy 2220 GCCCAGCCCTTAGAAACCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACG 2279
Db 2221 gccagccccctagaaaccccaagctctctctcggaaacccgctcacctagagccagac 2280
Qy 2280 TTACTAGGGTCTCTCCAGCTGTAGAGCTGAGGTTCACCCCTTAACCCAGGGAGCA 2339
Db 2281 ttaactcagggtcctccagctgttagagctgaggtttccaccttaacccaa-ggagca 2339
Qy 2340 CAGGTCCCACCTCCAGCCCGGGAGCCTAGGACCACTCAGCCCTTACGAGTATATTTCG 2399
Db 2340 caggtccacatccagccc-gggagccttaggacccactcagccccctagagtatattcc 2398
Qy 2400 CACTTCAGAAATTCATTTGCGAATCCAAGCTCCCTGCGCCCAAAATAACTTCAGTCTG 2459
Db 2399 cacttcagaaatccatctctggaatcccaagctccctgcccccaataacttcagtcct 2458
Qy 2460 CTTCCAGAAATTTGGAATTCCTAGTTTCTCTCTTCCTTCCTATCCCGAGTGTGGACACAAA 2519
Db 2459 c-tccagaatttgaaacctagtttctctctctctctctctctctctctctctctctct 2517
Qy 2520 CTCGCCCGCCAGCTATGAGCATCTTGAGCCCGCCCTCTTCTTCCAGCAAACTGGCCCG 2579
Db 2518 ctccgccccagctatgagcatctgagcccgccccctctctctgaagaaactggcccc 2577
Qy 2580 GATCAGAGCAGGAGCTCCTCTCCAGCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTC 2639
Db 2578 gatcagacagagacctccctccagacctctggaaacctccagaggtccagccatctc 2637
Qy 2640 GGAGCATCCCGAGGAATCTGACAGAGGGGTAGAGTGGGTGACAGAGCCTGATCTCT 2699
Db 2638 ggagcatcccgaggaaatctgcaga-gggtagagtgggtagacagagcgtgatctct 2696
Qy 2700 TCCCTTTTGTACATAGATTATTTTTCAGTTTCCAAAGAAAGATGAATATTTTGTAAA 2759

Db 2697 tctgtttttgacatagatttttttttcagttccaagaagaagatgaatcacattttgttaa 2756
Qy 2760 AAAAAAAAAA 2770
Db 2757 aaaaataaaa 2767
RESULT 4
US-60-278-561-1743
: Sequence 1743, Application US/60278561
: GENERAL INFORMATION:
: APPLICANT: Morris, MacDonald
: APPLICANT: Lal, Preeti
: APPLICANT: Diep, Dinh
: TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
: TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
: FILE OF INVENTION: Polymorphisms Identified Thereby
: FILE REFERENCE: GX-0012-1 P
: CURRENT APPLICATION NUMBER: US/60/278,561
: CURRENT FILING DATE: 2001-03-23
: NUMBER OF SEQ ID NOS: 15598
: SOFTWARE: PERL Program
: SEQ ID NO 1743
: LENGTH: 3534
: TYPE: DNA
: ORGANISM: Homo sapiens
: FEATURE:
: NAME/KEY: misc feature
: OTHER INFORMATION: Incyte ID No: 074428.3
US-60-278-561-1743

Query Match 93.1%; Score 2580.2; DB 8; Length 3534;
Best Local Similarity 98.1%; Pred. No. 0;
Matches 2706; Conservative 0; Mismatches 43; Indels 10; Gaps 9;
Qy 2 GTGTCCGAGCGCGGACCTTAGATCCAGAACCCAGCAGCGCAGCGCGCCGCGCA 61
Db 782 ggcacctagggcggaacttcatgtccacagaggtccggcgcggtgcgcgcgggtgc 841
Qy 62 CTATTTCCAGGCTCTGCGGGGAGGCGCGCGCCCGCCCGCCCGCCCGCATGGGCTGAA 121
Db 842 ctagt--gtggagatgtaagctcgatgccgcgccacccgccccgcctcagctggag 899
Qy 122 GCGCGCCAGAGAGCGTGTTCGCGCTGCGCTCCATCGACAGCGTGTGCGCTCTTTGC 181
Db -900 ggcgcgccagaaagcgtgttcccgctgcgtccatcagcagcgtggcgccgtgttc 959
Qy 182 TGCGAGCTGGCGCGCAGAGGAGCGGACCTGGTCTCTCTTTCTTTGCTGGGCTTCGT 241
Db 960 tgcgagctggcgagagagcgagcgagcctgtgtctcttctctgtgtgtggcctct 1019
Qy 242 GGAGCATTTCTGCTGTCAACCGCTCATCCCTAGCAACGTTCCCGAGCTCACCTTCCA 301
Db 1020 ggaacattttctgtgtcaaccgctcatctccctaccacgcttcccagagctcacctcca 1079
Qy 302 GCCCAGCGCCCGCCCGCAGCGCTTCCGCGCTCACCTTAC-TTTCCCGTGGCGACCTGT 360
Db 1080 gccacgc 1139
Qy 361 CTATCATCGCGCGCTCTATGCGCGCTTACCGCGCAGATCCGAGCGCGCGCTGACCTGT 420
Db 1140 ctatcatcgccgcctctatgcccgccttccgcgcgcgcgcgcgcgcgcgcgcgcgc 1199
Qy 421 CCGCTATCTCTCGAGAGGGGTGTCTCCAGCGGTGAGCTGTGAAGAGGTCTCCGATG 480
Db 1200 cctctctatctcgagaaaggggtgtctctccagcgtgtgagctgtgtgaaagaggtctcc 1259
Qy 481 TCATATGSAACAGCCTCAGCGCTCTCTACTTCAAGAGTTCGGGCGCCACATCCAGTCCCTCT 540
Db 1260 tcatatggaacagcctcagcgcctctacttcaagagtcggggccacatccagtcctct 1319
Qy 541 TCAGCTTCATCACAGGCCACCAATTTGGACAGCTCCGGTGTGGCCCTTTTGTGTGGG 600

|||||
Db 1320 tcaagttcatacaggacacaaattggaagactccgggtgtggcccttctgtgtggtggg 1379
QY 601 CTTGCCAGGCCCTGGTCTCCGGGATGTCCACCTCGCCCTGTCTGAGGATCATGCTCGG 660
Db 1380 cctgcgaagccctgggtctccgggtgtccacctcgccctgtcctgagatcatgctggg 1439
QY 661 TAGTGTGTGGCCCAATGGGAGCAGACAGCTGAGGTGACCTTGGCAGCGCAAGGCAAG 720
Db 1440 tagtgtttgggcccattgggagagacagctgaggtcaccttggcacggaaggaacag 1499
QY 721 AGGACCGAGGGCCAGACAGTCAATGCGGTGTGGCTGAGCGGAGCT--GGTGTACCTG 779
Db 1500 aggaaccgagggccagacagtcacatgcgggtgtgctgagcgagctgggtgtacctg 1559
QY 780 AAAGGATCATACATCGCTGTGACCGCAAGATGGAGGTGGCGTTCATGTGTGTGCCATC 839
Db 1560 aaagatcatacatcgtctgtgaccgcaaga tggggtgtcctcatgggtgtgtgcca tc 1619
QY 840 AACCCCTTCATTGACCTGCACACCGACTCGCTGGAGCTTCTGAGCTGCAGCAGAGCTG 899
Db 1620 aaccttcattgacctgcacacgactcgtctggagcttctgagctgcagcagaagctg 1679
QY 900 CTCGTGCTCTATGACTGGGACATCTGGAAGGTATCCCATGCCCTTAGGGAACTCG 959
Db 1680 ctctggtgctctatgacctgggacatctgaaaggtacccatggccttagggaaacctg 1739
QY 960 GCAGATCTAGAGAGCTGGAGCCACCCCTGCGCGCCAGACCACTACCCCTTACCAC 1019
Db 1740 gcagatctagagagctggagccccccctggcgccagacacacacacacacacacac 1799
QY 1020 AAGGCATTGCTCAGCCAAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTAC 1079
Db 1800 aaggcattgctcagcaagacctactatcggtatgaacacatctacccctacatgtac 1859
QY 1080 CTGGCTGCTACACTGTGCGCAACCGCAATGTGCGGGAAGCCCTGAGGCCCTGGCGGAC 1139
Db 1860 ctggtggtacacctgtgcgaaccgcaatgtgcgggaagccctgcagcctgggcggac 1919
QY 1140 ACGGCCACTGTCTACAGGACTACAACTACTCGCGGAGACGAGAGATCTACAGGAG 1199
Db 1920 acggccactgtctatccaggactcacactactgcccgggaagagagagatctcacaggag 1979
QY 1200 TTCTTTGAAGTAGCCCAATGATGTATCCCAACCTGCTGAAGGAGCAGCCAGCTGTCTG 1259
Db 1980 ttctttgaagtagccaatgattcatcccaacctgtgaagaggagcaagcagcttctg 2039
QY 1260 GAGGCGGCGAGAGCGCGCGGGAGCAAAAGCCAGGCAAGCCAGCCAGCAAGGTTCCGCC 1319
Db 2040 gaggcgggcagagcgcgcggggagcaaaagccagggcacccagagccaaaggttccgpc 2099
QY 1320 CTCCAGGACCTGAGTGTCTGCCCCACCTGCTGCGATTTCTAGACGGCATCTGCAATGG 1379
Db 2100 ctccaggacctgtgtctgcgccacctgtctgcattctgaagggcatctgcaaatgg 2159
QY 1380 GAGGAGGCGAGTCCACGCTGTGTGACGTGGGTGGCGCCACCTTTCTTGTGAGTCC 1439
Db 2160 gaggaggcagctcccaagcctgtgtgaagtgggctgggcccaccttcttgtgaagtc 2219
QY 1440 CTAGGCGGTTTGGAGGACAGTGGCGGAGAGAGGTGCGCATAGTAGCCGAGAGCGCCGAG 1499
Db 2220 ctaggcggttttgaggagacaggtgcgcaagaggtgcgcatagtgagcgagagcgcgag 2279
QY 1500 GGGCGGCGGCGAGGAGCGTGGCGGAGGAAGCCCGGAAGGCCGGCGCGGGGCCCA 1559
Db 2280 gcggcgagggccgagagcggtgggagagagagccgggaagggcgcgggggccca 2339
QY 1560 CGGCGGAGCTCCAAGCCAGAGGAGCCCCCGCCGCAAGAACGACAGCTGGACAGGCG 1619
Db 2340 cggcggaagtcacagccagagagagcccccgccgcccacgaagacagcactggacaaggcg 2399
QY 1620 CTGGGCAACCGGCGAGGCTGAGTGTGAGA--CCCCCGCGGAAGCCCTCTGGGAGCTGTCC 1678
|||||

Db 2400 ctgggacccggccagggtgtagtgtcaggaccccccccggaagcctctctggtacltgc 2459
QY 1679 TGGCAGAGCCCGAGGCGCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCGACATC 1738
Db 2460 tggcacagcccgagccctgaagg tggcagcagcgctcagg tgcacagcccgcaagca tc 2519
QY 1739 ACCACCGCGGAGGTTCCAGTGTCTCATTCTCAGAGTGAAGATGAAGGGCATGAAGGA 1798
Db 2520 accacgcgcggaggtccagtgctcaacttccagagtgagaagatgaaggcgcatgaagga 2579
QY 1799 GCTGTGTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGCACATC 1858
Db 2580 gctgtgtgtggccaccaagatcaactcgagcgccatcaagctgcaactcacgacagtc 2639
QY 1859 GCAAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGTGACTACACTCTGTCTTCT 1918
Db 2640 gcaagtgcagatgaagaagcagaagtg tccaccctctagtgtactacactctgtcttctc 2699
QY 1919 CAAGCGGCGAGCCAAAGGCTCTGAACTACTGGGACTTCTGGACCGCTTGTGGGGACCCA 1978
Db 2700 caagcggaagcgaagcgctctgaactactgggacttcggacgcgttgtgggaccca 2759
QY 1979 GSGTCG--CCTTAGTGTCCCAACTCTGAGCCCATGTGTCTGCCCCAGCCCAAGGGGACA 2037
Db 2760 ggctcgcccttagtcccccaactctgagcccatgtctgcccccgacccaaggagca 2819
QY 2038 GGCCTCACCTCTAGCCAAACCCCTAGGTTCGGTCCCGAGTACAGTCTGTATCAAAACCCA 2097
Db 2820 ggctcacctctacccaacccctaggttccccggtcccggagtacag tctgtatacaaccca 2879
QY 2098 CGATTTCCTCCAGCTCAGAAACCCAGGGCTCTGCCCGAGTCTGTAGAAATATAGTCTCTTC 2157
Db 2880 cgattttctccagctcagaaccagggtctctgccccag tctgtagaatataaggtctcttc 2939
QY 2158 TCCAGATATCCAGCGCGCCCAATGGAAACCTCAGCGTGGGTCTTAATTACCAGTCTTTAA 2217
Db 2940 tcccaagatcccgagcgccaatggaaacctcacgctgggtcctaataattaccagctcttaa 2999
QY 2218 AGSCCGAGCCCTAGAAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCCA 2277
Db 3000 agggccagccctagaaacccaagctctctctctcggaaccgctcaacctagagccagaccaa 3059
QY 2278 CGTTACTCAGGGCTCTCTCCAGCTTGTAGAGCTGAGGTTTACCCTTAACCCAAAGGAG 2337
Db 3060 cgttactcaggtctctccagcttgttagaggtttcaccccttaacccaa--ggag 3118
QY 2338 CACAGTCCCACTCCAGCGCGGGAGCCTAGGACCACTCAGCCCTAGAGTATATTTC 2397
Db 3119 cacaggtcccaacctccagccc--gggagcctaggaacctcagccccctagagtatatttc 3177
QY 2398 CGCACTTCAGAAATCCATATCTTGGCAATCCCAAGCTCCCTGCCCAAAATAACTTCAGTCT 2457
Db 3178 cgcacttcgaatctccatctcttgcaatcccaagctccccctgcacccaaataaccttcagtc 3237
QY 2458 TGTCTCCAGAAATTTGAAATCTTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2517
Db 3238 tgc-tccagaatttgaaatctctgttctctctctctgtatccccgag tcttgggacaaa 3296
QY 2518 AACTCGGCGCCGAGCCCTATGAGCATCTCAGCGCCCGCCCTCTTCTCTGACGAACCTGGGCC 2577
Db 3297 aactcgcccccgagcctatgagcatctgagccccctctctctctctctctctctctctctct 3356
QY 2578 CGGATCAGAGCAGGACCTCCCTTCCGACCTCTGGGAACCTCCACAGGTCAGGCCCATC 2637
Db 3357 cggatcagagcaggaacctccccctctcggaacctctcggaacctccccagag tccagccccatc 3416
QY 2638 TGGGAGCTCCCGGAGGAAATCTGCAGAGGGTTAGAGTGGGTGAGTGGGTGAGAGGCTGATCT 2697
Db 3417 tcggagatcccgaggaataatctgcaga--gggttagagtggtgacaaagagcctgatct 3475
QY 2698 CTTCTCTTTTGTACATAGATTTATTTTTCAGTTCACGAAGATGAATACATTTGTT 2756
Db 3476 ctctcgttttgtacatagatt 3534

RESULT 5

US-09-758-466-258
; Sequence 258, Application US/09758466
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM036
; CURRENT APPLICATION NUMBER: US/09/758,466
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 814
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 258
; LENGTH: 2168
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-758-466-258

Query Match 73.8%; Score 2046.6; DB 6; Length 2168;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 2112; Conservative 3; Mismatches 2; Indels 6; Gaps 6;

QY	651	CATGCCCTGGGTAGTGTGGGCCCCAATGGGAGCAGACAGCTGAGGTCACTGGCAGGCG	710
DB	32	caagcctggtagtgtttgngcccaatggggagcagacagtgaagtcacctggcagcgc	91
QY	711	AAGGCAACAGAGACCCAGCGGGCCACACAGTCAATCCCGGTGTGGCTGAGCGAGCTGG	770
DB	92	aagggaacagagaccgacgggcccagacagtcacatcccggtgtgctgagcggagcgtg	151
QY	771	CTGTACTGAAAGGATCATACATCGCTGTGTACCGCAAGATGAGGTGGGTTCATGGTG	830
DB	152	ctgtacctgaagatcatacatgcgtgtgacccgaagatggaggtggcggttcattgtg	211
QY	831	TGTGCCATCAACCTTCCATTGACCTGCACACGACFCGCTGGAGCTTCGACGCTCAG	890
DB	212	tgfgccatcaaccttccattgacctgcacacgacctgcgtggagcttctgagcgtgcag	271
QY	891	CAGAAGCTGCTGTGGCTGCTATGACCTGGGACATCTGGAAGAGTACCCCATGGCCTTA	950
DB	272	cagaagctgctgtgctgtctatgacctgggacatctggaaggttaccccatggcctta	331
QY	951	GGNAACCTGGCAGATCTAGAGAGCTGGAGCCACCCCTGGCGGGCAGACCCACTCACC	1010
DB	332	gggaacctggcagatctagaggagctggagcccacctggccggccagacccactcacc	391
QY	1011	CTCTACCAAGGCAATGTGCTCAGCCAAGACCTACTATCGGATGAACATCTACCCC	1070
DB	392	ctctaccacaaggagcattg-ctcagccaagacctactatcctggatgacacatctacccc	450
QY	1071	TACATGTACCTGGCTGCTACCACTGTGCAACCGCAATGTGGGGGAAGCCCTGCAGGCC	1130
DB	451	tacatgtacctggctggtgtaccactgtcgcaaccgcaatgtgggggaagccctgcagggc	510
QY	1131	TGGCGCAGACGGCCACTGTCTATCCAGGACTACACTACTGTCCGGGGAAGCAGGAGATC	1190
DB	511	tgggcgacacggccactgtcattccaggactacaactactgtccggggaagcaggagatc	570
QY	1191	TACAAGGAGTCTTTGAAGTAGCCCAATGATGTATCCCAACCTGCTGAAGGAGGACGCC	1250
DB	571	tacaaggagtctcttgaagtagccaa tga tgcattcccaacctgtcgaaggaggagcc	630
QY	1251	AGCTTGTGGAGCGGCGGAGGAGCGCCCGGGGAGCAAAAGCCAGGCGACCCACGACCAA	1310
DB	631	agcttgtgraggcgggcagagcgccggggagcaagccagggcagcccccagagccaa	690
QY	1311	GGTTCGCCCTCCAGGACCCCTGAGTGTCTCGCCCACTGTGCGATTCCTAGGACGGCATC	1370

DB	691	ggttccgcctccagaccctgagtgcttgcgccacctgctgcgattctacga tggcacc	750
QY	1371	TGCAATGGAGGAGGAGGAGGAGTCCACAGCCCTGTGCTGCACGTGGGCTGGCCACCTTTC	1430
DB	751	tgcaaatgggagggagggcagtcgccacgctgtgctgcacgtgggctgggacaccttctt	810
QY	1431	GTCCAGTCCCTAGGCGCTTTTGGAGGACAGGTGCGGAGAGGTGCGCATATGTCAGCCGA	1490
DB	811	gtccagtcccttagggccgttttgaggagacaggtgcgcagaaaggtgcgcalatg tga	870
QY	1491	GAGCCGAGGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1550
DB	871	gaggcgagcgccgagggcgagggagccgtggggcgagagcccgagggcgagggcgag	930
QY	1551	CGGGGCCCGGCGGAGTCCAAAGCCAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1610
DB	931	cggggccacggcgggaggtccaaagcagagggagcccccgcgcgccaagagagcagc	990
QY	1611	GACAAGGGCCTGGGCGACCGGCGAGGCTGCAGTGTGACGAGCCCGCCCGGAGGCTTC	1670
DB	991	gacaaggcgctgggcaccggccaggggtgcaggtgcagagacccccccgaaagcctc	1050
QY	1671	ACTGTGCTGGCAGACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1730
DB	1051	actgtcgctggcagcgccgagggccctgaaggtggcagcagggcctcaggtgcgagc	1110
QY	1731	GCAGCATCACACCGCGGAGGAGTCCAGTGTCTCAGTTCCTCAGAGTGAGAGTGAAGG	1790
DB	1111	gcagcatcacccacgcgcggagggtccaggtcctcactctccagagtgagagaga tga	1170
QY	1791	ATGAAGGAGTGTGTGGTGGCCACCAAGATCAACTCGAGCGCCCATCAAGCTGCAACT	1850
DB	1171	atgaaggagctggtgtggccacaaagatcaactcgagcgccatcaagctgcgaactc	1230
QY	1851	GCAGTGTGCAAGTGCAGATGAAGAGCAGCAAAAGTGTCCACCCCTAGTACTACACT	1910
DB	1231	gcacagtgcgaagtgcagatggaagagcagaaggtgtccacccctagtgactacactc	1290
QY	1911	TCCTTTCCTCAAGCGGAGCGCAAGGCGCTGTAACCTACTGGGACTTCGGACCGCTT	1970
DB	1291	tcttctctcaaggcgagcgagcgmaaaagcctctgaactactctggggaactcggagc	1350
QY	1971	GGAGCCAGGCTCCG-CCCTTAGTCCCGCAACTCTGAGCGCCATGTTCGCCCGCCAG	2029
DB	1351	gggacccaggtccgccttagtcccccaactctgagcccatgttctgccccagcgccaa	1410
QY	2030	AGGGACAGGCTCAGCTTACCCAAACCCCTAGGTTCCCGTCCCGAGTACAGTCTCTAT	2089
DB	1411	agggagcagggctcacctctacccaaacctaggttccccggtccccgagtagcagtc	1470
QY	2090	CAAAACCCAGCTTCTCCAGCTCAGAACCCAGGCGCTCTGCCCGAGTCGTTAGAAAT	2149
DB	1471	caaaaccaagatttctcagctcagaacccagggctctgccccagtcggtagaata tag	1530
QY	2150	GTCTTCTTCCAGAAATCCAGCGGCGCAATGGAACCTCACGCTGGGTCTTAATACCA	2209
DB	1531	gtcttcttccagaaatcccgagcgccaatggaaacctcagcgtgggttcttaattacca	1590
QY	2210	GTCTTTAAAGGCGCGCCCTAGAACCCAGCTCTCTCCGGAACCGCTACCTTACGC	2269
DB	1591	gtctttaaggccagcccttagaaacccagctctctctcggaaacccgctcaacctag	1650
QY	2270	CAGACCAACCTTACTCAGGGCTCTCTCCAGCTCTGTAGGAGCTGAGGTTTACCCCT	2329
DB	1651	cagaccaagtactcagggctctctccagcttctgaagagctgaaggttccaccttaacc	1710
QY	2330	CAGGGAGCAGAGTCCCACTCCAGCCCGGGAGCGCTAGACACACTCAGCCCTTAGGAG	2389
DB	1711	caa-ggagcagaggtcccaacctccagccc-gggagcctagagccactcagccccctag	1768
QY	2390	TATATTCCCGCACTTCAGAAATTCATATCTGCAATTCAGAGTCCCTCCGCCCCAAAT	2449

Db 1769 tatatttcgcgaattccatattatcttgcgaatccaagctccctccgcccccaataac 1828
QY 2450 TTCAGTCCTGTCCTCCAGAAATTTGAAATCTAGTTTCCTCTCTCGTATCCCGAGTCTG 2509
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1829 ttccagtcctgc-tccagaatttggaatcctctagtttccctctcttcgtatcccgagtcgt 1887
QY 2510 GCACACAAACTCCCGCCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAA 2569
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1888 ggacacaaactccgccccagcctatgacatcctgagcccgccctctctctgacgaa 1947
QY 2570 ACTGGCCCCGGATCAGAGCAGCCTCCCTCCGACCCCTCTGGGAACCTCCCGAGAGTCC 2629
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1948 actggccccgagatcagagcagactccctccgacctctgggaacctcccgagaggtcc 2007
QY 2630 AGCCCATCTCGAGCATCCCGGAGGAATCTGCAGAGGGTTAGGAGTGGGTGACAAAGAG 2689
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 2008 agcccatctcgagcatcccgaggagaaatctgcaga-gggttaggagtggtgacaaagag 2066
QY 2690 CCGTATCTCTCTCTGTTTGTACATAGATTTATTTTCAGTTCACGAAAGATGAATACA 2749
Db 2067 cctgatctctctctgtttgtacatagatttttttcagttcccaagaagaatgaataca 2126
QY 2750 TTTTGTAAAAAATAAAAAA 2772
Db 2127 tttgttaaaaaaataaaaaa 2149

RESULT 6

US-09-380-337-3
; Sequence 3, Application US/09380337
; GENERAL INFORMATION:
; APPLICANT: Chandrasekharappa, Setara C.
; Guru, Siradanahalli C.
; Manickam, Pachaiappan
; Collins, Francis S.
; Emmert-Buck, Michael R.
; Debelenko, Larisa V.
; Lubensky, Irina A.
; Liotta, Lance A.
; Agarwal, Sunita K.
; Spiegel, Allen M.
; TITLE OF INVENTION: MEN1, the Gene Associated with Multiple
; Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses
; Thereof
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380,337
; FILING DATE: 09-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/040,269
; FILING DATE: 05-MAR-1997
; APPLICATION NUMBER: WO PCT/US98/04258
; FILING DATE: 04-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Lockyer, Jean M.
; REGISTRATION NUMBER: 44,879
; REFERENCE/DOCKET NUMBER: 015280-315100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300

; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 9180 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: -
; LOCATION: 1..9180
; OTHER INFORMATION: /note= "genomic sequence for MEN1 gene"
; FEATURE:
; NAME/KEY: exon
; LOCATION: 1680..1766
; FEATURE:
; NAME/KEY: intron
; LOCATION: 1767..2264
; FEATURE:
; NAME/KEY: exon
; LOCATION: 2265..2732
; FEATURE:
; NAME/KEY: intron
; LOCATION: 2733..4296
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4297..4505
; FEATURE:
; NAME/KEY: intron
; LOCATION: 4506..4715
; FEATURE:
; NAME/KEY: exon
; LOCATION: 4716..4844
; FEATURE:
; NAME/KEY: intron
; LOCATION: 4845..5176
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5177..5217
; FEATURE:
; NAME/KEY: intron
; LOCATION: 5218..5297
; FEATURE:
; NAME/KEY: exon
; LOCATION: 5298..5385
; FEATURE:
; NAME/KEY: intron
; LOCATION: 5386..6024
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6025..6161
; FEATURE:
; NAME/KEY: intron
; LOCATION: 6162..6622
; FEATURE:
; NAME/KEY: exon
; LOCATION: 6623..6758
; FEATURE:
; NAME/KEY: intron
; LOCATION: 6759..7195
; FEATURE:
; NAME/KEY: exon
; LOCATION: 7196..7360
; FEATURE:
; NAME/KEY: intron
; LOCATION: 7361..7577
; FEATURE:
; NAME/KEY: exon
; LOCATION: 7578..8881
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-380-337-3

Query Match

47.3%; Score 1312.2; DB 5; Length 9180;

Best Local Similarity 96.5%; Pred. No. 1.4e-245;
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1382 GGAGGCGAGTCCACAGCCTGCTGCTGACGCTGGGCTGGCCACCTTTCTTCTGTCAGTCCCT 1441
Db 7499 GGATGGCGAGCAGAGGCTCTGGAGTTCCAGCACCTGCGCGGAACCTTCTCTCTCACTT 7558
QY 1442 AGGCCGTTTGGAGGACAGGTGGCGGAGAGGTGCGCATAGTACGCGAGAGCCGAGGC 1501
Db 7559 GCTCTCCCACTGGCCAGGTGGCGGAGAGGTGCGCATAGTACGCGAGAGCCGAGGC 7618
QY 1502 GGCGAGGCGAGAGCGGTGGGCGAGAGAGCCCGGGAAGCGGCGGCGGCGGCCACG 1561
Db 7619 GGCGAGGCGAGAGCGGTGGGCGAGAGAGCCCGGGAAGCGGCGGCGGCGGCCACG 7678
QY 1562 GCAGGAGTCCAGGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1621
Db 7679 GCGGAGTCCAGGCGAGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7738
QY 1622 GGGCAGCGGCGAGGTCAGTGTGAGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1681
Db 7739 GGGCAGCGGCGAGGTCAGTGTGAGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7798
QY 1682 CACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1741
Db 7799 CACAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7858
QY 1742 ACCCGCGGAGGTCAGTGTGAGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1801
Db 7859 ACCCGCGGAGGTCAGTGTGAGGACCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7918
QY 1802 GCTGGTGGCCACCAAGATCAACTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1861
Db 7919 GCTGGTGGCCACCAAGATCAACTCGAGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 7978
QY 1862 AGTGCATGAGAGAGAGAGAGTGTCCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTC 1921
Db 7979 AGTGCATGAGAGAGAGAGTGTCCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGGTCAGG 8038
QY 1922 GCGCAGCGCAAGGCGCTCTGAAGTCTGGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGG 1981
Db 8039 GCGCAGCGCAAGGCGCTCTGAAGTCTGGGCGGCTGGGCGGCGGCGGCGGCGGCGGCGGCGG 8098
QY 1982 TCGCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCGCCCGGCGGCGGCGGCGGCGGCGGCGG 2041
Db 8099 TCGCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCGCCCGGCGGCGGCGGCGGCGGCGGCGG 8158
QY 2042 TCAGCTCTACCCAAACCTAGGTTCCCGGTCGCGGAGTACAGTCTGTATCAACCCAGCAT 2101
Db 8159 TCAGCTCTACCCAAACCTAGGTTCCCGGTCGCGGAGTACAGTCTGTATCAACCCAGCAT 8218
QY 2102 TTTCTCAGCTCAGAACCCAGGCTCTGCGCCAGTCTGTTAGATATAGTCTCTTCTCC 2161
Db 8219 TTTCTCAGCTCAGAACCCAGGCTCTGCGCCAGTCTGTTAGATATAGTCTCTTCTCC 8278
QY 2162 AGAATCCAGCGCGGCCAATGAAACCTCAGCGTGGGTCCTAATACAGTCTTTAAAGGC 2221
Db 8279 AGAATCCAGCGCGGCCAATGAAACCTCAGCGTGGGTCCTAATACAGTCTTTAAAGGC 8338
QY 2222 CCAGCCCTTAGAATCCCAAGCTCTCTCGAACCCTCAGCTAGAGCCAGACCAAGCTT 2281
Db 8339 CCAGCCCTTAGAATCCCAAGCTCTCTCGAACCCTCAGCTAGAGCCAGACCAAGCTT 8398
QY 2282 ACTCAGGCTCTCCAGCTCTGAGGAGTGTAGGTCCTTACCCCTTAACCCAGGAGGACACA 2341
Db 8399 ACTCAGGCTCTCCAGCTCTGAGGAGTGTAGGTCCTTACCCCTTAACCCAGGAGGACACA 8458
QY 2342 GGTCCCACTCCAGCGCGGAGCGCTAGGACCACTACGCGCCCTAGGAGTATATTTCCGCA 2401
Db 8459 GGTCCCACTCCAGCGCGGAGCGCTAGGACCACTACGCGCCCTAGGAGTATATTTCCGCA 8518
QY 2402 CTTCAGAAATCCCATATCTTGGGAATCAAGCTCCCTGCGCCCAATACCTTACGTCCTGCT 2461
|||||

Db 8519 CTTCAGAAATCCCATATCTTTCGGAATCCAAAGTCCCTGCCCAAAATACATTCAGTCCCTGCT 8578
QY 2462 TCCAGAAATTTGGAATCTCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2521
Db 8579 TCCAGAAATTTGGAATCTCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8638
QY 2522 CCGCCCGCCAGCCTATGAGCATCTCTGAGCCCGCCCTCTTCTCTGACGAACTGGCCCGCGGA 2581
Db 8639 CCGCCCGCCAGCCTATGAGCATCTCTGAGCCCGCCCTCTTCTCTGACGAACTGGCCCGCGGA 8698
QY 2582 TCAGAGCAGACCTCCCTCCGACCCCTCTGGGAACCTCCAGAGGTCAGCCCATCTCGG 2641
Db 8699 TCAGAGCAGACCTCCCTCCGACCCCTCTGGGAACCTCCAGAGGTCAGCCCATCTCGG 8758
QY 2642 AGCATCCCGAGGAAATCTCAGAGGCGTTAGGAGTGGTGACAAGAGCCTGATCTCTTC 2701
Db 8759 AGCATCCCGAGGAAATCTCAGAGGCGTTAGGAGTGGTGACAAGAGCCTGATCTCTTC 8818
QY 2702 CTGTTTGTACATAGATTTATTTTTCAGTTCCCAAGAAAGATGAATACATTTTGTAAAAA 2761
Db 8819 CTGTTTGTACATAGATTTATTTTTCAGTTCCCAAGAAAGATGAATACATTTTGTAAAAA 8878
QY 2762 AAAAAAAA 2770
Db 8879 AATATAAA 8887
||| | |||

RESULT 7
US-09-880-107-3461
; Sequence 3461, Application US/09880107
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3461
; LENGTH: 9180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. U93237
; NAME/KEY: unsure
; LOCATION: (1), (9180)
; OTHER INFORMATION: n - a o r c o r g o t
US-09-880-107-3461

Query Match 47.3%; Score 1312.2; DB 7; Length 9180;
Best Local Similarity 96.5%; Pred. No. 1.4e-245;
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 1382 GGAGGCGAGTCCACAGCCTGCTGCTGACGCTGGGCTGGCCACCTTTCTTCTGTCAGTCCCT 1441
Db 7499 GGATGGCGAGCAGGCTCTGGAGTTCCAGCACCTGCGCGGAACCTTCTCTCTCACTT 7558
QY 1442 AGGCCGTTTGGAGGACAGGTGGCGGAGAGGTGCGCATAGTACGCGAGAGCCGAGGC 1501
Db 7559 GCTCTCCCACTGGCCAGGTGCGGAGAGGTGCGCATAGTACGCGAGAGCCGAGGC 7618
QY 1502 GGCGAGGCGGAGGCGGTCGGGCGAGAGAGCCCGGGAAGCGGCGGCGGCGGCCACG 1561
Db 7619 GCGGAGGCGGAGGCGGTCGGGCGAGAGAGCCCGGGAAGCGGCGGCGGCGGCCACG 7678
|||||

Db 846 GCGGCGCGCAAGGCGCTCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCAGGC 787
Qy 1982 TCCG-CCTTAGTCCCACTCTGAGCCCATGTTCTGCCCCCAAGGAGGACAGGC 2040
Db 786 TCCGCGCTTAGTCCCACTCTGAGCCCATGTTCTGCCCCCAAGGAGGACAGGC 727
Qy 2041 CTCACCTCTACCAACCCCTAGGTTCCGGTCCGGAGTACAGTCTGTATCAAAACCCACGA 2100
Db 726 CTCACCTCTACCAACCCCTAGGTTCCGGTCCGGAGTACAGTCTGTATCAAAACCCACGA 667
Qy 2101 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCACTGTTAGATATAGTCTCTCTCC 2160
Db 666 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCACTGTTAGATATAGTCTCTCTCC 607
Qy 2161 CAGAAATCCAGCGCGGCAATGGAACCTCAGCTGGGTCTTAATTACAGCTTTTAAAGG 2220
Db 606 CAGAAATCCAGCGCGGCAATGGAACCTCAGCTGGGTCTTAATTACAGCTTTTAAAGG 547
Qy 2221 CCCAGCGCTTAGAAACCCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGT 2280
Db 546 CCCAGCGCTTAGAAACCCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGT 487
Qy 2281 TACTCAGGCTCTCCAGCTTGTAGAGCTGAGGTTTCAACCTTAACCCAAAGGAGCAC 2340
Db 486 TACTCAGGCTCTCCAGCTTGTAGAGCTGAGGTTTCAACCTTAACCCAAAGGAGCAC 428
Qy 2341 AGGTCCACCTTCCAGCGCGGAGCTTAGGACCACTCAGCCCTTAGGATATATTTCGGC 2400
Db 427 AGGTCCACCTTCCAGCGCGGAGCTTAGGACCACTCAGCCCTTAGGATATATTTCGGC 369
Qy 2401 ACTTCAGAAATTCATATCTTGGCAATCCAGCTCCCTGCCCCCAATTAACCTCAGTCTCG 2460
Db 368 ACTTCAGAAATTCATATCTTGGCAATCCAGCTCCCTGCCCCCAATTAACCTCAGTCTCG 309
Qy 2461 TTCAGAAATTTGGAATTCCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2520
Db 308 -TTCAGAAATTTGGAATTCCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 250
Qy 2521 TCCGCGCGCGAGCTATGAGCATCTGAGCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCT 2580
Db 249 TCCGCGCGCGAGCTATGAGCATCTGAGCCCGGCTCTCTCTCTCTCTCTCTCTCTCTCT 190
Qy 2581 ATCAGACGAGACCT 2640
Db 189 ATCAGACGAGACCT 130
Qy 2641 GAGCATCCGAGGAAATCTCAGAGGGTTAGGAGTGGGTGACAAGAGCTGATCTCTT 2700
Db 129 GAGCATCCGAGGAAATCTCAGAGGGTTAGGAGTGGGTGACAAGAGCTGATCTCTT 71
Qy 2701 CCTGTTTGTACATAGATTTATTTTTCAGTTCCAAAGAGATGAATACATTTTGTAAAA 2760
Db 70 CCTGTTTGTACATAGATTTATTTTTCAGTTCCAAAGAGATGAATACATTTTGTAAAA 11
Qy 2761 AAAAAAATAA 2770
Db 10 AAAAAAATAA 1

RESULT 9
US-09-948-941-529
; Sequence 529, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948, 941
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0

; SEQ ID NO 529
; LENGTH: 11190
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-529

Query Match 45.2%; Score 1253.2; DB 5; Length 11190;
Best Local Similarity 96.2%; Pred. No. 4.3e-234;
Matches 1337; Conservative 0; Mismatches 48; Indels 5; Gaps 5;

Qy 1382 GGAGGGCAGTCCACAGCCTGTGTGCACGTGGGTGGGCGCCACCTTTCTTGTGCAGTCCCT 1441
Db 7819 ggaatggcagcagaggtccttgaggttccagccactggccggaacacttctctcaactt 7878
Qy 1442 AGGCCGTTTGTGGGACAGGTGGCGGAGAGGTGCGCATAGTGAAGCCGAGAGCCGAGGC 1501
Db 7879 gctctcccactggcccaggtgcygaggaaggtgcygcatagtgaagccgagagggc 7938
Qy 1502 GGCGGAGCGCGAGAGCCGTGGGCGAGAGCCCGGAAAGCGCGGCGGCGGCCACG 1561
Db 7939 ggcgagggccgagagcgctg9ggcgaggaagcccg9gag9ccg9c9gg9ccacg 7998
Qy 1562 GCGGAGTCCAAAGCCAGAGGAGCCCGCGCGCCCAAGAACGACGACTGGACAGGCGCT 1621
Db 7999 gcygaggtccaaagcagagagggcccgccgccaagaaagccagcactggacaagggc 8058
Qy 1622 GGCACCGCGCAGGTGTCAGTGTGAGGACCCCGCGGAAGCCTCTCTGGGACTGTGCTGG 1681
Db 8059 gggcaccg9gcaag9gtgagtgcaag9ccccc9gag9ccctctctg9gactg9c9ctg 8118
Qy 1682 CACAGCGCGAGGCGCTGAAGGTGGCAGCAGCGGTCTAGGTGCCAGCCCGCAGCATCACC 1741
Db 8119 cacagcccgagggccctgaaggtggcagcagcagctcagtgccagcagccagcagcacc 8178
Qy 1742 ACGCCCGAGGTCAGTGTCTACTTTCCAGAGTGAAGATGAAGGGCATGAAGAGAGCT 1801
Db 8179 acg9cgaggggtccagtgctcacttccagagtgagaagatgaagggcatgaagggagct 8238
Qy 1802 GCTGTGGCGCACCAAGATCAACTGAGCGCATCAAGCTGCAAGCTCAGCGCACAGTCCGA 1861
Db 8239 gcyg9gtggccacgaatcaactcagcagcagcagcagcagcagcagcagcagcagcagc 8298
Qy 1862 AGTGCAGATGAAGAGAGAGGTGTCCACCCCTAGTACTACTACTCTGTCTCTCTCTCTCT 1921
Db 8299 agtgagatgaagagcagaaagtgtccacccctagtgactacactctgctcttctctcaa 8358
Qy 1922 GCGCAGCGCAAGGCGCTCTCAACTACTGGGAGTTCGGACCGCTTGGGAGCCAGGC 1981
Db 8359 gcyg9cag9caagggcctctgaactactgg9gactctg9gagc9ctctg9gg9cccg9gc 8418
Qy 1982 TCCG-CCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGC 2040
Db 8419 tcg9ccctagtcccccacactctgagcccatgttctgccccagcccaagggagcagc 8478
Qy 2041 CTCACCTCTACCAACCCCTAGGTTCCGGTCCGAGTACAGTCTGTATCAAAACCCACGA 2100
Db 8479 ctcaactctaccacaaacccctaggttccg9gtcccg9gtacagctgctatcaaacccacga 8538
Qy 2101 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCGAGTCTGTTAGATATAGTCTCTCTCTCC 2160
Db 8539 ttttctccagctcagaaacccagggctctgccccagctcgttgaagataag9ctctctctcc 8598
Qy 2161 CAGAAATCCAGCGCGCAATGGAACCTCAGCTGGGTCTCTTAATTACAGCTTTTAAAGG 2220
Db 8599 cagaatcccgccg9ccaaatggaacactcagcgtg9gtctctaattaccagctcttaag9 8658
Qy 2221 CCCAGCGCTTAGAAACCCAGGCTCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGT 2280
Db 8659 cccagcccttagaaacccagcctctctctcgtggaaccgctcactcagtagagccagcagcact 8718
Qy 2281 TACTCAGGCTCTCCAGCTTGTAGGAGCTGTGAGGAGTTCACCTTAAACCCAGGAGCAC 2340

[illegible]

```

RESULT 10
US-09-948-941-649
; Sequence 649, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948,941
; CURRENT FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 649
; LENGTH: 11205
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-649

Query Match          45.2%; Score 1253.2; DB 5; Length 11205;
Best Local Similarity 96.2%; Pred. No. 4.3e-234;
Matches 1337; Conservative 0; Mismatches 48; Indels 5; Gaps

QY 1382 GGAGGGCAGTCCACGCGCTGTCTCCAGCTGGGCTGGGCCACCTTTCTTGTGCGAGTCCT 1441
    ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7819 ggaatgccagagcagggtctctgagttccagccactggccggcaacctgtctctcaacct 7878

QY 1442 AGCGCGTTTTTGAGGGACAGGTGCGGCAGAAAGTGCGCATAGTGAGCCGAGAGGCCGAGGC 1501
    ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7879 gctctccccacttgccccagggtgcgcagaaagtgcgcataatgagccgagagcccgaggc 7938

QY 1502 GGCCGAGGCGGAGGAGCCGTGGGGCGGAGAAAGCCGGGAAGCGCGGGCGGGGCCCCACG 1561
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 7939 ggcagagccgagagcagtcgtgggcgcagaaagcccggaagccgcgcgcgcgcgcgcgc 7998

QY 1562 GCGGGAGTCCAAAGCCAGAGGAGGCCGCCGCCCAAGAACGACGACTGGACAAAGGCGCT 1621
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```


Query Match 17.1%; Score 473.4; DB 6; Length 519;
Best Local Similarity 99.0%; Pred. No. 9e-83;

RESULT 14

;
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 2000-09-21
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: US 09/774,203
; PRIOR FILING DATE: 2001-01-29
; NUMBER OF SEQ ID NOS: 49117
; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
; SEQ ID NO 5337
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000134.14
; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
; US-09-864-761-5337

Query Match 16.3%; Score 452; DB 7; Length 464;

Best Local Similarity 99.8%; Pred. No. 1.3e-78; Indels 1; Gaps 1;
Matches 463; Conservative 0; Mismatches 0;

QY 1597 AGAAGCAGCACTGGACAAAGGGCTGGGACCGGAGGCTGAGGTGTCAGGACCCCGCC 1656
DB 464 AGAAGCAGCACTGGACAAAGGGCTGGGACCGGAGGCTGAGGTGTCAGGACCCCGCC 405
QY 1657 GGAAGCCTCTGGGACTGTGCTGGGACACCGCGGAGGCTGAGGTGCGAGCAGGCTC 1716
DB 404 GGAAGCCTCTGGGACTGTGCTGGGACACCGCGGAGGCTGAGGTGCGAGCAGGCTC 345
QY 1717 AGTGCCAGCACCGCAGCATCACACCGCGGAGGCTGAGGTGCTCACTTTCCAGAGTG 1776
DB 344 AGTGCCAGCACCGCAGCATCACACCGCGGAGGCTGAGGTGCTCACTTTCCAGAGTG 285
QY 1777 AGAAGATGAAGGGCATGAAGAGAGTGTGTGGCCACCAAGATCAACTCGAGGCGCATCA 1836
DB 284 AGAAGATGAAGGGCATGAAGAGAGTGTGTGGCCACCAAGATCAACTCGAGGCGCATCA 225
QY 1837 AGCTGCAACTCAGGCGACAGTCGCAAGTGCAGATGAAGAGCAAGAGTGTCCACCCCTA 1896
DB 224 AGCTGCAACTCAGGCGACAGTCGCAAGTGCAGATGAAGAGCAAGAGTGTCCACCCCTA 165
QY 1897 GTGACTACACTCTCTCTTCTCAAGCGGAGCGCAAGGCTCTGAACCTACTTGGGACT 1956
DB 164 GTGACTACACTCTCTCTTCTCAAGCGGAGCGCAAGGCTCTGAACCTACTTGGGACT 105
QY 1957 TCGGACCGCTTGTGGGACCCAGGCTCCG-CCCTAGTCCCGCCCACTCTGAGCCCATGTC 2015
DB 104 TCGGACCGCTTGTGGGACCCAGGCTCCGCTTGTCCCGCCCACTCTGAGCCCATGTC 45
QY 2016 TGGCCCCAGCCCAAGGGGAGGAGCGCTCACTCTACCCCAACCC 2059
DB 44 TGGCCCCAGCCCAAGGGGAGGAGCGCTCACTCTACCCCAACCC 1

RESULT 15
US-60-278-561-1743/c

; Sequence 1743, Application US/60278561
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0012-1 P
; CURRENT APPLICATION NUMBER: US/60/278,561
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 1743
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 074428.3
; US-60-278-561-1743

Query Match 16.1%; Score 447; DB 8; Length 3534;
Best Local Similarity 96.8%; Pred. No. 1.7e-77;
Matches 456; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 86 GGGCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCC 145
DB 719 GGGCG 660
QY 146 GCTGCGCTCCATCGACGACGTGGTGGCCCTGTTTGTGCGCGAGCTGGCGCGAGAGGAC 205
DB 659 GCTGCGCTCCATCGACGACGTGGTGGCCCTGTTTGTGCGCGAGCTGGCGCGAGAGGAC 600
QY 206 GGACCTGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 265
DB 599 GGACCTGGTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 540
QY 266 CGTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGCCCGCCCGCCCGCCCGCCCGCC 325
DB 539 CGTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGCCCGCCCGCCCGCCCGCCCGCC 480
QY 326 TGGGGGCTCACCTACTTTCCTGCGCGGACCTGTCTATCATCGCCCGCTCTATGCGCGG 385
DB 479 TGGGGGCTCACCTACTTTCCTGCGCGGACCTGTCTATCATCGCCCGCTCTATGCGCGG 420
QY 386 CTTACCGCGCCAGATCGAGGCGCGCTGACCTGTCCCTCTATCTCGAAGAGGGGTGT 445
DB 419 CTTACCGCGCCAGATCGAGGCGCGCTGACCTGTCCCTCTATCTCGAAGAGGGGTGT 360
QY 446 CTCAGCGCTGAGCTGGTGAAGAGGTCTCCGATGTCTATGGAACAGCTTCAGCCGCTC 505
DB 359 CTCAGCGCTGAGCTGGTGAAGAGGTCTCCGATGTCTATGGAACAGCTTCAGCCGCTC 300
QY 506 CTACTTCAAGGATCGGGCGCCACATCCAGTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCT 556
DB 299 CTACTTCAAGGATCGGGCGCCACATCCAGTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCT 249

Search completed: October 14, 2001, 08:17:47
Job time: 16272 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 21:11:14 ; Search time 13115.4 Seconds
(without alignments)
3123.151 Million cell updates/sec

Title: US-09-380-337-1

Perfect score: 2772

Sequence: 1 GGTGTCGGAGCCGCGGACC.....TGTTAAAAA.....2772

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 14155048 seqs, 7388405095 residues

Total number of hits satisfying chosen parameters: 28310096

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main:*

1: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US08_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US099_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US0001_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US0002_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US0003_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US0004_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US0005_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US0006_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US0007_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US0008_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US0009_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US0010_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US0011_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US0012_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US0013_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US0014_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US0015_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US0016_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US0017_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US0018_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US0019_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US0020_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US0021_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US0022_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US0023_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US0024_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US0025_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US0026_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US0027_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US0028_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2587.2	93.3	2840	49	US-60-172-360-22297
2	1865.2	67.3	2196	12	US-08-865-337-2
3	863.8	31.2	986	12	US-08-865-337-4
4	791.8	28.6	32768	53	US-60-213-178-196
5	746.4	26.9	812	12	US-08-865-337-5
6	522.2	18.8	603	20	US-09-539-800-12611
7	492.2	17.8	538	12	US-08-865-337-3
8	488.4	17.6	639	1	PCR-US01-00663-18748
9	486.2	17.5	551	21	US-09-540-229-133603
10	471.6	17.0	917	29	US-09-721-589-5669
11	470	17.0	867	29	US-09-721-589-4604
12	459.2	16.6	465	19	US-09-528-409-62595
13	452	16.3	464	1	PCR-US01-00663-5551
14	448	16.2	784	29	US-09-726-211-1570
15	446	16.1	2121	22	US-09-577-408-6141
16	446	16.1	2770	22	US-09-577-408-136
17	440	15.9	479	16	US-09-205-070-29321
18	440	15.9	479	17	US-09-340-623-29321
19	440	15.9	479	17	US-09-353-690-11417
20	436.4	15.7	549	27	US-09-698-013-349
21	425	15.3	461	27	US-09-698-010-4746
22	399	14.4	443	28	US-09-716-990-1013
23	393	14.2	405	17	US-09-332-782-18044
24	393	14.2	405	19	US-09-515-694-18044
25	378	13.6	406	19	US-09-528-409-92976
26	360.8	13.0	513	51	US-60-196-710-1772
27	356.6	12.9	381	24	US-09-637-086A-46608
28	356.6	12.9	381	25	US-09-654-617-229277
29	356.6	12.9	381	27	US-09-684-016-229277
30	323	11.7	447	17	US-09-306-350A-32588
31	318	11.5	391	16	US-09-248-797-47016
32	318	11.5	391	17	US-09-346-956-12130
33	310	11.2	389	16	US-09-235-076-11698
34	310	11.2	389	16	US-09-248-797-37577
35	310	11.2	389	17	US-09-332-782-11698
36	310	11.2	389	29	US-09-737-223-11698
37	308.4	11.1	2442	49	US-60-171-625-524
38	308.4	11.1	2442	49	US-60-173-464-25060
39	308.4	11.1	2442	51	US-60-191-637-32160
40	308.4	11.1	2442	51	US-60-191-681-25485
41	308.4	11.1	2442	53	US-60-219-005-305
42	295.6	10.7	413	12	US-08-803-609-4307
43	287.8	10.4	305	14	US-09-044-767-468
44	287.8	10.4	305	21	US-09-540-764-23525
45	279.8	10.1	310	21	US-09-540-212A-61207

ALIGNMENTS

```

RESULT      1
US-60-172-360-22297
; Sequence 2297, Application US/60172360
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide Polymor
; FILE REFERENCE: GX-0007 P
; CURRENT APPLICATION NUMBER: US/60/172,360
; CURRENT FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 29838
; SOFTWARE: PERL Program
; SEQ ID NO 2297
; LENGTH: 2840
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 074428.3
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2836
; OTHER INFORMATION: a, t, c, g, or other
US-60-172-360-22297

```

Query Match 93.3%; Score 2587.2; DB 49; Length 2840;
Best Local Similarity 98.0%; Pred. NO. 0;
Matches 2713; Conservative 0; Mismatches 44; Indels 10; Gaps 9;

Qy	2	GTGTCCGGAGCCGCGGACCTAGAGATCCGAGAGCCACAGCGAGCGGCCCGCCGCGCA	61
Db	79	gcgacctagggcggaactctatgtcccgacggtccgcggctgcgcgcgggtgc	138
Qy	62	CTATTTCCAGGCTCTCGGGGAGGGCGCGCCACCGCCGCGCCATGGGCTGAA	121
Db	139	ctagt--gtggagtgtaaagtctggaTgcgcgcgcacccgcgcgcctatgggtgaa	196
Qy	122	GGCGCCACAGACGGTGTCCCGTGGGCTCCATCGACGAGTGGTGGGCTGTTTC	181
Db	197	ggccgccagaagaagctgttccgctgcgtcccaTcgaagacgtggtgcgctattgc	256
Qy	182	TGGCGAGCTGGCGGAGAGAGCGGACCTGTGTCTTTCTTGGTGGCTGGGCTTGT	241
Db	257	tgcgagctggcgagagagcggaactggtgtcccttctcttggtgctgggttgt	316
Qy	242	GGACATTTTCTGGCTCTCAACCGCTCATCCCTACCAAGTTCCCGAGCTCACCTTCCA	301
Db	317	ggagcatcttctggctgtcaacgcgcgtcatccctacaaagttcccgagctcaccttcca	376
Qy	302	GCCAGCGCCGCCCCGACCGCTGGCGGCTCACCTACTTTCCCGTGGCGAGACTGTC	361
Db	377	gcccagcccgcgcccgcgacccgcgtggcgctctacttctcccgTggcgacctgtc	436
Qy	362	TATCATCGCGCCCTCTATGCCCGCTTCACGCCCCAGATCCGAGGCGCGCTCGACTGTC	421
Db	437	talcatcgcgccctctaTgcccgtctcaccgcccagaTccgaggcgcgctgcacctgtc	496
Qy	422	CCCTATCTTCGAGAGAGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAGGTCTCGAGTGT	481
Db	497	ccctctaTccctcgagaagggtgtctccagcgtgagctggTgaagaaggtctccgagt	556
Qy	482	CATATGAACAGCTCAGCGGCTCTACTTCAAGAGTACGGGCCACATCCAGTGCCTCTTT	541
Db	557	caTatggaaagcctcagcgcTccTacttcaaggatcgggcccaTccagTccctctt	616
Qy	542	CAGCTTCATCACAGGACCCAAATTGGACAGCTCCCGTGTGGCCTTTGTGTGGTGGGC	601

[illegible]

QY 268 TCATCCCTACCAACGTTCCCC

APPLICANT: Tang, Y. Tom
APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
TITLE OF INVENTION: PROTEIN
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,337
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 986 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT08
CLONE: 2809030
US-08-865-337-4

Query Match 31.2%; Score 863.8; DB 12; Length 986;
Best Local Similarity 92.9%; Pred. No. 2.4e-167;
Matches 907; Conservative 2; Mismatches 63; Indels 4; Gaps 2;
QY 668 TGGGCCCAATGGGGAGCAGACAGCTGAGTGCACCTGGCAGCGCAAGGGCAAGGACCG 727
DB 1 TGGGCCCAATGGGGAGCAGACAGCTGAGTGCACCTGGCAGCGCAAGGGCAAGGACCG 60
QY 728 CAGGGCCAGACAGTCAATCGCGGTGGCTGAGCGGAGCTGGCTGTACCTAAAGGATC 787
DB 61 CAGGGCCAGACAGTCAATCGCGGTGGCTGAGCGGAGCTGGCTGTACCTAAAGGATC 120
QY 788 ATACATGGCTGTACCGCAAGATGGAGTGGCTTTCATGGTGTGTCGCATCAACCCCTTC 847
DB 121 ATACATGGCTGTACCGCAAGATGGAGTGGCTTTCATGGTGTGTCGCATCAACCCCTTC 180
QY 848 CATTGACCTGCACACCGACTCGCTGGAGTTCTGAGCTGCAGAGCAAGAGCTCTGTGGCT 907
DB 181 CATTGACCTGCACACCGACTCGCTGGAGTTCTGAGCTGCAGAGCAAGAGCTCTGTGGCT 240
QY 908 GCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCTTGGGAACCTGGCAGATCT 967
DB 241 GCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCTTGGGAACCTGGCAGATCT 300
QY 968 AGAGGAGCTGGAGCCACCCCTGGCGGCGCAGACCCACTCACCTCTACCAAGGGCAT 1027
DB 301 AGAGGAGCTGGAGCCACCCCTGGCGGCGCAGACCCSRGGGCGCTCTACCAAGGGCAT 360
QY 1028 TGCCTCAGCCAAAGACCTACTATTCGGGATGAACACATCTACCCCTACATGTACCTGGCTGG 1087

DB 361 TGTTCAGCCCAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACCTGGCTGG 420
QY 1088 CTACCACCTGTGCGCAACCGCAATGTGCGGAAGCCCTCGAGGCTGGG---CGACACGGC 1144
DB 421 CTACCACCTGTGCGCAACCGCAATGTGCGGAAGCCCTCGAGGCTGGGCGAGAGACACGGC 480
QY 1145 CACTGTCTATCCAGGACTACAACTACTGCGGGGAAGCAGGAGAGATCTACAAGAGTTCTT 1204
DB 481 CACTGTCTATCCAGGACTACAACTACTGCGGGGAAGCAGGAGAGATCTACAAGAGTTCTT 540
QY 1205 TGAAGTAGCCCAATGATGTCATCCCAACCTGCTGAAGGAGGAGCCAGCTTCTGGAGGC 1264
DB 541 TGAAGTAGCCCAATGATGTCATCCCAACCTGCTGAAGGAGGAGCCAGCTTCTGGAGGC 600
QY 1265 GGGCGAGGAGCGCGCGGGGAGCAAGCCAGGCGACCCAGAGCCCAAGGTTCCGCCCTCCA 1324
DB 601 GGGCGAGGAGCGCGCGGGGAGCAAGCCAGGCGACCCAGAGCCCAAGGTTCCGCCCTCCA 660
QY 1325 GGACCCCTGAGTGTCTGCGCCACCTGCTGCGATTCTACGACGCATCTGCAAAATGGAGGA 1384
DB 661 GGACCCCTGAGTGTCTGCGCCACCTGCTGCGATTCTACGACGCATCTGCAAAATGGAGGA 720
QY 1385 GGGCAGTCCACCGCTGCTGCTGCGAGTGGGCTGGCCACCTTCTTGTGTCAGTCCCTAGG 1444
DB 721 GGGCAGTCCACCGCTGCTGCTGCGAGTGGGCTGGCCACCTTCTTGTGTCAGTCCCTAGG 780
QY 1445 CGGTTTGGAGGACAGGTGCGGCGACAAAGTGTGAGCCGAGAGCGGAGCGCGC 1504
DB 781 CGGTTTGGAGGACAGGTGCGGCGACAAAGTGTGAGCCGAGAGTGCAGAGATGCCGAGGCTGC 840
QY 1505 CGAGCCGAGGAGCGCTGGGCGAGGAAGCCGCGGGAAGCCGCGGCGCGCCACGCGG 1564
DB 841 CGAGGCC-ANGAGCCGTGGNGCTAGGAAGCCCGGNAANGCCGCGGCGCCACNGCA 899
QY 1565 GGAGTCCAAAGCCAGAGGAGCGCCCGCGGCCCAAGAGCCAGCAGCTGGACAAAGGCTGGG 1624
DB 900 GGAGTCCAAAGCCANATGAGCNCNCCGCGCNAANNAGCNANNAGCAGNAGNCCTTTGGG 959
QY 1625 CACCGCGCAGGCTGCA 1640
DB 960 CMC CGCCAGNGTGA 975

RESULT 4
US-60-213-178-196/C
; Sequence 196, Application US/60213178
; GENERAL INFORMATION:
; APPLICANT: Besley, Ellen
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000689
; CURRENT APPLICATION NUMBER: US/60/213,178
; CURRENT FILING DATE: 2000-06-22
; NUMBER OF SEQ ID NOS: 1425
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 196
; LENGTH: 32768
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)-(32768)
; OTHER INFORMATION: n = A,T,C or G
US-60-213-178-196

Query Match 28.6%; Score 791.8; DB 53; Length 32768;
Best Local Similarity 99.2%; Pred. No. 4.2e-152;
Matches 848; Conservative 0; Mismatches 2; Indels 5; Gaps 5;
QY 1917 CTCAAGCGCAGCGCAAGGCGCTCTGAACTACTGGGAGCTTCGAGCCGCTTGTGGGAGC 1976

Db 32768 CTCAAGCGGAGGCGCAAGAGGCTCTGAACCTACCTGGGACTTCGGACGCTTGTTGGGGACC 32709
QY 1977 CAGGCTCCG-CCTTAGTCCCGCCCACTAGAGCCATGTTCTGCGCCCGCCAGCCCAAGGGGA 2035
Db 32708 CAGGCTCCGCGCTTAGTCCCGCCCACTAGAGCCATGTTCTGCGCCCGCCAGCCCAAGGGGA 32649
QY 2036 CAGGCTCAGCTCTACCCAAACCTAGGTTCCCGGTCGCCAGTACAGTCTGTATCAAAACC 2095
Db 32648 CAGGCTCAGCTCTACCCAAACCTAGGTTCCCGGTCGCCAGTACAGTCTGTATCAAAACC 32589
QY 2096 CAGGATTTCTCAGCTCAGAACCCAGGCTCTGCGCCAGTCTGTAGATATAGTCTCT 2155
Db 32588 CAGGATTTCTCAGCTCAGAACCCAGGCTCTGCGCCAGTCTGTAGATATAGTCTCT 32529
QY 2156 TCTCCAGAAATCCAGCGGCGCAATGAAACCTCAGCTGGGTCTTAATACCAAGTCTTT 2215
Db 32528 TCTCCAGAAATCCAGCGGCGCAATGAAACCTCAGCTGGGTCTTAATACCAAGTCTTT 32469
QY 2216 AAAGGCCAGCCCTAGAAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACC 2275
Db 32468 AAAGGCCAGCCCTAGAAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACC 32409
QY 2276 AAGTTACTCAGGCTCTCTCCAGCTCTGAGAGCTGAGGTTTCACCTTAACCCAAAGG 2335
Db 32408 AAGTTACTCAGGCTCTCTCCAGCTCTGAGAGCTGAGGTTTCACCTTAACCCAAAGG 32350
QY 2336 AGCAGAGTCCCACTCCAGCGCGGAGCTAGGACCACTAGGACCTAGGAGTATATT 2395
Db 32349 AGCAGAGTCCCACTCCAGCGGAGCTAGGACCTAGGAGTATATT 32291
QY 2396 TCCGCACTTCAGAAATCCATATCTTGGAAATCCAGCTCCCTGCCCAATAACTTCAGT 2455
Db 32290 TCCGCACTTCAGAAATCCATATCTTGGAAATCCAGCTCCCTGCCCAATAACTTCAGT 32231
QY 2456 CTTGCTTCCAGAAATTTGGAATCCTAGTTCTCTCTGATCCGAGTCTGGGACAC 2515
Db 32230 CTTGCTTCCAGAAATTTGGAATCCTAGTTCTCTCTGATCCGAGTCTGGGACAC 32172
QY 2516 AAAACTCCGCGCCCGAGCTATGAGCATCTGAGCGCGCGCTCTCTGACGAAACTGGC 2575
Db 32171 AAAACTCCGCGCCCGAGCTATGAGCATCTGAGCGCGCGCTCTCTGACGAAACTGGC 32112
QY 2576 CCGGATCAGAGCAGGACCTCTCTCGAGCCCTCTGGGAACCTCCAGAGGTCAGGCCCA 2635
Db 32111 CCGGATCAGAGCAGGACCTCTCTCGAGCCCTCTGGGAACCTCCAGAGGTCAGGCCCA 32052
QY 2636 TCTCGAGCATCCGAGGAAATCTGAGAGGTTAGGAGTGGGTGACAGAGCTGAT 2695
Db 32051 TCTCGAGCATCCGAGGAAATCTGAGAGGTTAGGAGTGGGTGACAGAGCTGAT 31993
QY 2696 CTCTCTCTGTTTGTACATAGATTATTTTTCAGTTCCAGAAAGATGAATACATTTTGT 2755
Db 31992 CTCTCTCTGTTTGTACATAGATTATTTTTCAGTTCCAGAAAGATGAATACATTTTGT 31933
QY 2756 TAAAAAATAATAAA 2770
Db 31932 TAAAAAATAATAAA 31918

RESULT 5
US-08-865-337-5
; Sequence 5, Application US/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,337
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 812 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT02
CLONE: J313372
US-08-865-337-5

Query Match 26.9%; Score 746.4; DB 12; Length 812;
Best Local Similarity 97.2%; Pred. No. 3.2e-143;
Matches 772; Conservative 4; Mismatches 16; Indels 2; Gaps 2;
QY 1570 CCAAGCCAGAGGAGCCCCCGCCCAAGAGCCAGCAGCTGGACAGGGCTTGGGACCG 1629
Db 1 CCAAGCCAGAGGAGCCCCCGCCCAAGAGCCAGCAGCTGGACAGGGCTTGGGACCG 60
QY 1630 GCCAGGTCAGTGTCCAGGACCCCGGAAAGCTCTCTGGGACTGTCGCTGGCAGACGCC 1689
Db 61 GCCAGGTCAGTGTCCAGGACCCCGGAAAGCTCTCTGGGACTGTCGCTGGCAGACGCC 120
QY 1690 GAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGACCCCGCAGCATCACCGCCGG 1749
Db 121 GAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGACCCCGCAGCATCACCGCCGG 180
QY 1750 AGGTCCTGCTCAGTCTCCAGAGTGAAGATGAAGGCGATGAAGGAGCTGCTGGTGG 1809
Db 181 AGGTCCTGCTCAGTCTCCAGAGTGAAGATGAAGGCGATGAAGGAGCTGCTGGTGG 240
QY 1810 CCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGCAGCAGTCCCAAGTGCAGA 1869
Db 241 CCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGCAGCAGTCCCAAGTGCAGA 300
QY 1870 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTACTCTGTCTTTTCTTCTTCAAGCGCAGC 1929
Db 301 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTACTCTGTCTTTTCTTCTTCAAGCGCAGC 360
QY 1930 GCAAGGCCCTCTGAAGTGTCCAGCCCTAGTGTGGGAGCTTGGGAGCCAGGCTCCG-CCT 1988
Db 361 GCAAGGCCCTCTGAAGTGTCCAGCCCTAGTGTGGGAGCTTGGGAGCCAGGCTCCG-CCT 420
QY 1989 TAGTCCCGCCCAACTCTGAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCTCCTACCTC 2048
Db 421 TAGTCCCGCCCAACTCTGAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCTCCTACCTC 480

```

QY 2049 TACCCAAACCCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 2108
Db 481 TACCCAAACCCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 540
QY 2109 AGCTCAGAACCCAGGCTCTGCCCGAGTCCGTTAGATATAGTCTCTTCTCCAGATCC 2168
Db 541 AGCTCAGAACCCAGGCTCTGCCCGAGTCCGTTAGATATAGTCTCTTCTCCAGATCC 600
QY 2169 CAGCGCGCCCAATGGAAACCTCAGCGTGGGTCTTAATTTACAGTCTTTAAAGGCCAGGCC 2228
Db 601 CAGCGCGCAATGGAAACCTCAGCGTGGGTCTTAATTTACAGTCTTTAAAGGCCAGGCC 660
QY 2229 CTAGAACCCCAAGCTCTCTCTCGG-AACCGCTCAGCTAGAGCCAGACCAACGTTACTCAG 2287
Db 661 CTAGAACCCCAAGCTCTCTCTCGGAAACCGTCACTAGAGCCAGACCAACGTTACTCAG 720
QY 2288 GGCTCTCCCGAGCTTGTAGAGCTGAGGTTTACCCCTTAACCCAGGGAGCAGCTGCC 2347
Db 721 GGCTCTCCCGAGCTTGTAGAGCTGAGGTTTACCCCTTAACCCAGGGAGGAAANGTCCC 780
QY 2348 ACTCCAGCCCGGG 2361
Db 781 ANCTTCAGCCCGGG 794

```

RESULT 6

```

US-09-539-800-12611
; Sequence 12611, Application US/09539800
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Deleageane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF CONNECTIVE TISSUE
; FILE REFERENCE: PD-1023 CIP
; CURRENT APPLICATION NUMBER: US/09539,800
; CURRENT FILING DATE: 2000-03-30
; PRIOR APPLICATION NUMBER:
; PRIOR FILING DATE: August 16, 1995
; PRIOR FILING DATE: June 27, 1994
; PRIOR FILING DATE: August 16, 1995
; PRIOR FILING DATE: June 27, 1994
; PRIOR FILING DATE: November 4, 1994
; PRIOR FILING DATE: November 4, 1994
; PRIOR FILING DATE: October 3, 1997
; PRIOR FILING DATE: October 3, 1997
; PRIOR FILING DATE: October 4, 1996
; PRIOR FILING DATE: October 4, 1997
; PRIOR FILING DATE: October 4, 1997
; PRIOR FILING DATE: October 4, 1996
; PRIOR FILING DATE: December 18, 1997
; PRIOR FILING DATE: December 18, 1997
; PRIOR FILING DATE: December 20, 1996
; PRIOR FILING DATE: December 20, 1996
; PRIOR FILING DATE: February 10, 1999
; PRIOR FILING DATE: February 10, 1999
; PRIOR FILING DATE: February 12, 1998
; PRIOR FILING DATE: February 12, 1998
; PRIOR FILING DATE: December 1, 1999
; PRIOR FILING DATE: December 1, 1999
; PRIOR FILING DATE: December 10, 1998
; PRIOR FILING DATE: December 10, 1998
; NUMBER OF SEQ ID NOS: 19698
; SOFTWARE: PERL Program
; SEQ ID NO 12611
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature

```

```

; OTHER INFORMATION: Incyte ID No: hu01347785
US-09-539-800-12611
Query Match 18.8%; Score 522.2; DB 20; Length 603;
Best Local Similarity 98.1%; Pred. No. 3.9e-97;
Matches 560; Conservative 0; Mismatches 8; Indels 3; Gaps 3;
QY 2038 GCCTCCTACCTCTACCCAAACCCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCCA 2097
Db 1 ggcctcacctctacccaaacccctaggtctcccggtcccgagtagcagtcgtatataagtcctctc 60
QY 2098 CGATTTCCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGTCCGTTAGATATAGTCTCTTTC 2157
Db 61 cgattttctccagctcagaacccagggctctgccccagtcgttagaataatagtcctctc 120
QY 2158 TCCCAAGATCCCGAGCCGCGCAATGGAAACCTCAGCGTGGGTCTTAATTTACCAAGTCTTTAA 2217
Db 121 tcccaagaatcccgagcgcgcgaatggaaacctcacgcgtgggtcctaattaccagtccttaa 180
QY 2218 AGCCCCAGCCCTTAGAAACCCAGGCTCTCTCGGAAACCGCTCACCTAGAGCCAGACCAA 2277
Db 181 aggccagccccctagaaacccagctctctctcggaacccagctcacctagagccagacccaa 240
QY 2278 CGTTACTCAGGGCTCTCCAGCTTGTAGAGCTGAGGTTTACCCCTTAACCCAAAGGAG 2337
Db 241 cgttactcagggctctccagctgttagagctgaggtttcaccccttaacccaa-ggag 299
QY 2338 CACAGTCTCCAGCTCAGCCGCGGGAGCCTTAGGACCACTCAGCCCTTAGGAGTATATTTTC 2397
Db 300 cacaggtccacctccagccc-gggagcctagggaccactcagcccttaggagtagtatattc 358
QY 2398 CGCAGCTTCAGAAATTCATATCTTGGCAATCCAAAGCTCCCTGCCCAAAATAACTTCAGTCC 2457
Db 359 cgcacttcagaaatccatattctgcgaatcccaagctccctgcgccccaaataacttcagtc 418
QY 2458 TGCTCCAGAAATTTGAAATTCCTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2517
Db 419 tgc-tccagaatttgaaatccatagttctctctctctctctctctctctctctctctctct 477
QY 2518 AACTCCGCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2577
Db 478 aactccgccccagcctatgagcatctgagcccgccctctctctctctctctctctctctctct 537
QY 2578 CGGATCAGAGCAGGAGCTCCCTTCGACCCCT 2608
Db 538 cggatcagagcagggactcccttcggacctct 568

```

RESULT 7

```

US-08-865-337-3
; Sequence 3, Application US/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; STREET: Incyte Pharmaceuticals, Inc.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/865,337
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 538 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT14
CLONE: 1596949
US-08-865-337-3

Query Match 17.8%; Score 492.2; DB 12; Length 538;
Best Local Similarity 99.2%; Pred. No. 5.7e-91;
Matches 502; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
QY 88 GCCGCGGCCACCGCCGCCCATGGGCTGAAGGCGGCCAGAACAGCGTGTTCGCCG 147
DB 34 GCCGCGGCCACCGCCGCCCATGGGCTGAAGGCGGCCAGAACAGCGTGTTCGCCG 93
QY 148 TGCCTCATCGACGAGTGTCGCGCTGTTGCTGCGGAGCTGGCGCGGAGAGGCGCG 207
DB 94 TGCCTCATCGACGAGTGTCGCGCTGTTGCTGCGGAGCTGGCGCGGAGAGGCGCG 153
QY 208 ACTGGTGCTCTTTTCCCTTGGTGTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACGCG 267
DB 154 ACTGGTGCTCTTTTCCCTTGGTGTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACGCG 213
QY 268 TCATCCCTACCAAGTTCGCGAGCTACCTTCAGGCCAGCCGCCGCCCGACCCGCGCTG 327
DB 214 TCATCCCTACCAAGTTCGCGAGCTACCTTCAGGCCAGCCGCCGCCCGACCCGCGCTG 273
QY 328 CGCGCTCACCCTACTTTCGCGTGGCGGACCTGTCTATCATCGCGCCCTCTATCGCGCT 387
DB 274 CGCGCTCACCCTACTTTCGCGTGGCGGACCTGTCTATCATCGCGCCCTCTATCGCGCT 333
QY 388 TCACCGCCCGAGATCCGAGGCGCGCTGACCTGTCCCTCTATCTTCGAGAAGGGGGTGTCT 447
DB 334 TCACCGCCCGAGATCCGAGGCGCGCTGACCTGTCCCTCTATCTTCGAGAAGGGGGTGTCT 393
QY 448 CCAGCGGTGAGCTGGTAGAGAGTCTCCGATGTCAATGGAACAGCTCAGCGGCTCCT 507
DB 394 CCAGCGGTGAGCTGGTAGAGAGTCTCCGATGTCAATGGAACAGCTCAGCGGCTCCT 453
QY 508 ACTTCAAGGATCGGGCCACATCCAGTCCCTCTTCTCAGCTTCATCAGAGGACCAAAATGG 567
DB 454 ACTTCAAGGATCGGGCCACATCCAGTCCCTCTTCTCAGCTTCATCAGAGGACCAAAATGG 513
QY 568 ACAGCTCGGTGTGGCCTTTGCTGTG 593
DB 514 ACA-STCGGNGTGGCCTTTGCTGTG 538

RESULT 8
PCT-US01-00663-18748/c
Sequence 18748, Application PC/TUS0100663
GENERAL INFORMATION:
APPLICANT: Molecular Dynamics, Inc.
APPLICANT: Penn, Sharon G.

APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
FILE REFERENCE: PB 0004 WO 7 PCT/US01/00663
CURRENT APPLICATION NUMBER: PCT/US01/00663
CURRENT FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 04 February 2000 (04.02.00)
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 26 May 2000 (26.05.00)
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 03 August 2000 (03.08.00)
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 03 October 2000 (03.10.00)
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 27 September 2000 (27.09.00)
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 21 September 2000 (21.09.00)
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 30 June 2000 (30.06.00)
NUMBER OF SEQ ID NOS: 38837
SOFTWARE: Molecular Dynamics Sequence Listing Engine
SEQ ID NO 18748
LENGTH: 639
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC000134.14
OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
OTHER INFORMATION: EST_HUMAN HIT: BF526328.1, EVALUATE 0.00e+00
OTHER INFORMATION: NT HIT: U93237.1, EVALUATE 0.00e+00
OTHER INFORMATION: SWISSPROT HIT: O00255, EVALUATE 2.00e-43
PCT-US01-00663-18748

Query Match 17.6%; Score 488.4; DB 1; Length 639;
Best Local Similarity 91.8%; Pred. No. 3.6e-90;
Matches 516; Conservative 0; Mismatches 46; Indels 0; Gaps 0;
QY 1382 GGAGGGGAGTCCACAGCGCTGTGTGACGTGGGTGGGCGCCACCTTTCTTGTGTGAGTCCCT 1441
DB 562 GGATGGCGAGAGGAGGTCTGTGGAGTTCCAGCCACTGCGCGGCAACCTTCTCTACCTT 503
QY 1442 AGGCGGTTTTTGAGGGACAGGTGCGGAGAGGTGCGCATAGTACGCGAGAGCGCGAGGC 1501
DB 502 GCTCTCCCCACTGGCCAGGTGCGGAGAGGTGCGCATAGTACGCGAGAGCGCGAGGC 443
QY 1502 GGCGGAGCGGAGAGCGGTGGGCGGAGAGCCCGGGAAGCGCGGCGGCGGCGGCGCCACG 1561
DB 442 GGCGGAGCGGAGAGCGGTGGGCGGAGAGCCCGGGAAGCGCGGCGGCGGCGGCGCCACG 383
QY 1562 GCGGAGTCCAGCCACAGGAGCGCCCGCGCCCAAGAGCCAGCTGACAGAGGCGCT 1621
DB 382 GCGGAGTCCAGCCACAGGAGCGCCCGCGCCCAAGAGCCAGCTGACAGAGGCGCT 323
QY 1622 GGGCAGCGGCGAGGTGTCAGTGTGAGGACCCCGCGGAGCGCTCTTGGGACTGTTCGCTGG 1681
DB 322 GGGCAGCGGCGAGGTGTCAGTGTGAGGACCCCGCGGAGCGCTCTTGGGACTGTTCGCTGG 263
QY 1682 CACAGCCCGAGGCGCTGAAGGTGGCAGCAGCGGTCTAGGTGCCAGCAGCCCGGAGCATCACC 1741
DB 262 CACAGCCCGAGGCGCTGAAGGTGGCAGCAGCGGTCTAGGTGCCAGCAGCCCGGAGCATCACC 203
QY 1742 ACCGCGGAGGTTCCAGTGTCTACTTTCCAGAGTGAAGATGAAGGCGCATGAAGGAGCT 1801
DB 202 ACCGCGGAGGTTCCAGTGTCTACTTTCCAGAGTGAAGATGAAGGCGCATGAAGGAGCT 143
QY 1802 GCTGGTGGCCACCAAGATCAACTGAGCGGCATCAAGCTGCAACTCAGGCGACAGTGCACA 1861
DB 142 GCTGGTGGCCACCAAGATCAACTGAGCGGCATCAAGCTGCAACTCAGGCGACAGTGCACA 83

```

Qy 1862 AGTCAGATGAAGAGCAGAGAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTTTCCTCAA 1921
Db 82 AGTCAGATGAAGAGCAGAGAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTTTCCTCAA 23
Qy 1922 GCGCAGCGCAAGAGCCCTCTGA 1943
Db 22 GCGCAGCGCAAGAGCCCTCTGA 1
RESULT 9
US-09-540-229-133603
; Sequence 133603, Application US/09540229
; GENERAL INFORMATION:
; APPLICANT: Seilhamer, Jeffrey J.
; APPLICANT: Delegeane, Angelo M.
; APPLICANT: Stuart, Susan G.
; APPLICANT: Stuve, Laura L.
; APPLICANT: Mullahy, Sara J.
; APPLICANT: Naughton, Rebecca E.
; TITLE OF INVENTION: POLYNUCLEOTIDES OF NERVOUS SYSTEM AND SENSORY ORGANS
; FILE REFERENCE: PD-1033 CIP
; CURRENT APPLICATION NUMBER: US/09/540,229
; CURRENT FILING DATE: 2000-03-31
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 193582
; SOFTWARE: PERL Program
; SEQ ID NO 133603
; LENGTH: 551
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: incyte ID No: hu01071364
; NAME/KEY: unsure
; LOCATION: 389, 423, 466, 484
; OTHER INFORMATION: a, t, c, g, or other
US-09-540-229-133603

```

```

Db 417 cccagcncatgagcabcctgagccgcgcctcttctgacgaaactgnccccgatcag 476
Qy 2586 AGCAGGAGACCTCCCTTCCGAGCCCTCTCGGAACTCTCCGAGAGGTCCAGCCCATCTCGGAGCA 2645
Db 477 agcaggancctccctccgaccctctgggaacctccagaggLccagcccactctcgagca 536
Qy 2646 TCCCGGAGGAATCT 2660
Db 537 ttcggaggaaatct 551

RESULT 10
US-09-721-589-5669
; Sequence 5669, Application US/09721589
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Vileval, Jean-Luc
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2045-001
; CURRENT APPLICATION NUMBER: US/09/721,589
; CURRENT FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 60/167,380
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 7017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5669
; LENGTH: 917
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-721-589-5669

Query Match 17.0%; Score 471.6; DB 29; Length 917;
Best Local Similarity 98.1%; Pred. No. 1.1e-86;
Matches 477; Conservative 0; Mismatches 9; Indels 0; Gaps

```

Query Match	17.5%	Score 486.2;	DB 21;	Length 551;
Best Local Similarity	97.8%	Pred. No. 9;8e-90;		
Matches 543; Conservative	0;	Mismatches 7;	Indels 5;	Gaps

Qy	2107	CCAGCTCAGAACCCAGGGGCTCTGCCCCAGTCGTTAGAAATATAGGTCTCTCTCCAGAAAT	2166
Db	1	ccagctcagaaccaca-ggctctgccccagtcgttagaataatagtcctcttcccagaat	59
Qy	2167	CCCAGCGCGCCAATGGAAACCTCACGCTGGGTGCTCTTAATTACAGTCTTTAAAGGCCCCAGC	2226
Db	60	cccagcgcgcaatggaaacctcagctgggtcctaataccagctcttaaaagcccaagc	119
Qy	2227	CCCTAGAAACCCAGCTCCTCTCTGGGAACCGCTCACCTAGAGCCAGACCAAGTTACTCA	2286
Db	120	ccctagaaacccaagctcctctcctggaaaccgctcacctagagcagagaccaagttactca	179
Qy	2287	GGGCTCTCTCCAGCTGTGAGGAGCTGAGGTTTACCCCTTAACCCAAAGGAGCACAGGTCC	2346
Db	180	gggctctctccagctgttagagagctgaggtttcacctcttaacccaa-ggagcacaggtcc	238
Qy	2347	CACCTCCAGCCCGGGAGCCCTAGGACCACTCAGCCCCCTAGGAGTATATTTCCGCACTTCA	2406
Db	239	cacctccagccc-gggagcctaggaccactcagccctcaggagtatatattccgcacttca	297
Qy	2407	GAATTCCATATCTTCCGAATCCAAGCTCCTCTGCCCCCAATAACTTCAGTCTCTGCTTCAG	2466
Db	298	gaattcccatattgcgaattccaagctccctgcgcccaataaacttcagtcctgc-tccag	356
Qy	2467	AATTTGGAAATCTTAGTTTCTCTCTCTTTCGTATCCCGAGTCTGGGACAC-AAAACCTCGCG	2525
Db	357	aacttgaaaaacctagtttctctctctcggtantccggagctcgggacacaaaaactccgcg	416
Qy	2526	CCCCAGCCTATGAGCATCCTGAGCCGCCCTCTTCTCTGACGAAACTGGCCCCCGGATCAG	2585

	Query Match	17.0%	Score	471.6;	DB	29;	Length	917;
	Best Local Similarity	98.1%;	Pred.	No. 1.le-86;				
	Matches	477;	Conservative	0;	Mismatches	9;	Indels	0;
	Gaps							
Qy	431	TCGAGAAAGGGGTCTCCACGCCGTGAGCTGGTGGAAGAAGTCTCCGATCTCATATCGAA	A90					
Db		3 tgcaccacgctcgccgccagcgtgagctggggaagaaggctcttcyagtccata tggaa	62					
Qy	491	CAGCCTCAGCCGCTCTACTTTCAAGGATCGGGGCCACATCCACGTCCCTCTTCACGTTCAAT	550					
Db		63 cagcctcagcgctcttactccaagatcgggccacatccagtcctcttcagcttc	122					
Qy	551	CACAGGCACCNAATTGACAGCTCCCGGTGGGCTTTGCTGTGGTTGGGGCTTGCCAGGC	610					
Db		123 cacaggcaccaattgacagctccgggtggcctttgctgtggttggtgggacctlgccagnc	182					
Qy	611	CCTGGGTCTCGGATGTCACACTCGCCCTGTGTGAGGATCATCGCTGGGTACTGTTTTGG	670					
Db		183 cctgggtctcgggatgtccacctgcctctgtctgaggatcata tgcctgggtagtgtlctgg	242					
Qy	671	GCCCAATGGGAGCAGACAGCTGAGGTACCTGSCACGGCAAGGGCAACGAGGACCCGAG	730					
Db		243 gcccaa tggggagcagacagctgagg tccactggcagcgcaagg gcaacgaggaaccgcag	302					
Qy	731	GGGCCACACAGCTCAATGCCGGTGTGGCTGAGCGAGCTGGCTTACCTGAAAAGCATATA	790					
Db		303 ggcccagacagtcaatgccggtgtggctgagcagagctggcttacctgacatgaagatcata	362					
Qy	791	CATCGCTGTGACCCGAAGATGGAGGTGGCGTTTCATGGTGTGGCCATCAACCTCTCCAT	850					
Db		363 catcgctgtgaccgcgaagatggagg tggcggtcata tgggtgtgtgcacataaaccttccat	422					
Qy	851	TGACCTGCACACCGACTCGCTGGAGCTTCTGCAGTGCACGAGAAGTGCTCTGGCTGTGT	910					
Db		423 tgacctgcacacegactcgtctggagctcttgcagctgcagcagaagctgctcttgggtgct	482					
Qy	911	CTATGA	916					
Db		483 ctatga	488					

RESULT 11

US-09-721-589-4604
; Sequence 4604, Application US/09721589
; GENERAL INFORMATION:
; APPLICANT: Gearing, David P.
; APPLICANT: Villevail, Jean-Luc
; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 1600.2045-001
; CURRENT APPLICATION NUMBER: US/09/721,589
; PRIOR FILING DATE: 2000-11-22
; PRIOR APPLICATION NUMBER: 60/167,380
; PRIOR FILING DATE: 1999-11-24
; NUMBER OF SEQ ID NOS: 7017
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4604
; LENGTH: 867
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: (1)...(867)
; OTHER INFORMATION: n = A,T,C or G
US-09-721-589-4604

Query Match 17.0%; Score 470; DB 29; Length 867;

Best Local Similarity 97.9%; Pred. No. 2.4e-86;
Matches 476; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

Qy	431	TCGAGAAGGGGGTCTCTCCAGCCGTGAGCTGGTGAAGAAGCTCTCCGATGTCATATGAA	490
Db	12	tcgcccacgcgtccgcctccagcctgagctggtgaagaaggtcccgatgcatatgaa	71
Qy	491	CAGCCTCAGCCGCTCCTACTTCACAGGATCGGGCCACATCCAGTCCCTCTTCACCTTCAT	550
Db	72	cagctcagcgcgtccctacttcaaggatcgggccacatccagtcctcttcagcttcat	131
Qy	551	CACAGGACCAAAATTGGACAGCTCCGGTGTGGCCTTTCTGCTGGGGCTGCCAGGC	610
Db	132	caaggcaacaaattggaacagctccggtgtggtcccttctgtgttggcctccaggc	191
Qy	611	CCTGGGTCTCCGGATGTCACCTCGCCCTGTCTGAGGATCATGCCGTGGGTAGTGTGG	670
Db	192	cctgggtctccggatgtccaccctcgccctgtctgagatcatgctggtgtgttgg	251
Qy	671	GCCCAATGGGGAGCAGACAGCTGAGGTACCTGGCAGGCAAGGCGCACGAGCCGAG	730
Db	252	gcccaatggggagcagacagctgaggttcacctggcagcgcaaggagacccgag	311
Qy	731	GGGCACAGACTCAATGCGGTGTGGGTGAGCGGAGCTGGCTGTACCTGAAAGGATCATA	790
Db	312	gggccagacagtcgaatcccggtgtggtgagcgagctggtgtacctgaaagatcata	371
Qy	791	CATGGCTGTGACCGCAAGATGAGGTGGCGCTTCATGGTGTGCCATCAACCCCTTCAT	850
Db	372	catgctgtgacccgaagatggaggtggcgcttcatggtgtgtgccatcaaccccttcat	431
Qy	851	TGACCTGCACACCCGACTCGCTGGAGCTTCTGCAGCTGCAGCAGAGCTGCTCGCTGCT	910
Db	432	tgacctgcacaccgactcgtcgtggaggttcttcagctgcagcagagctgctcctggtcgt	491
Qy	911	CTATGA 916	
Db	492	ctagga 497	

RESULT 12

US-09-528-409-62595
; Sequence 62595, Application US/09528409
; GENERAL INFORMATION:
; APPLICANT: Drmanac, Radoje T.

; APPLICANT: Labat, Ivan
; APPLICANT: Stache-Crain, Birgit
; APPLICANT: Dickson, Mark
; APPLICANT: Jones, Lee W.
; TITLE OF INVENTION: Novel Nucleic Acid Sequences Obtained
; FILE REFERENCE: 774
; CURRENT APPLICATION NUMBER: US/09/528,409
; CURRENT FILING DATE: 2000-03-17
; PRIOR APPLICATION NUMBER: 60/125,453
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 116231
; SOFTWARE: Hy-patent.pl Version 3.1
; SEQ ID NO 62595
; LENGTH: 465
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(465)
; OTHER INFORMATION: n = A,T,C or G
US-09-528-409-62595

Query Match 16.6%; Score 459.2; DB 19; Length 465;
Best Local Similarity 99.4%; Pred. No. 3.3e-84;
Matches 461; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy	638	CCTGCTGAGGATCATGCCCTGGGTAGTGTGTTGGCCCAATGGGAGCAGACAGCTGAGGT	697
Db	1	cctgctgagatcatcctcctgggtagttgttggcccaatgggagcagacagctgaggt	60
Qy	698	CACCTGGCAGCGCAAGGCAACGAGGACCGAGGGCCAGACAGTCAATGCCGTGTGGC	757
Db	61	cacctggcagcgcaaggcagggcgagggccagagggccagacagtcgaatgccggtgtggc	120
Qy	758	TGACGGAGCTGGCTGTACCTGAAAGGATCATATGCGCTGTGACCGCAAGATGGAGGT	817
Db	121	tgaagcgagctggctgtacctgaaagatcatatcgctgtgacggcaagatggaggt	180
Qy	818	GGCCTTCATGCTGTGTCATCAACCTTCCATTTGACCTGCACACCGACTCGCTGGAGCT	877
Db	181	ggccttcattggtgtgtccatcaaccttccattgacctgcacaccgactcgtggagct	240
Qy	878	TCTCAGCTGCAGCAGAGCTGCTCTGGCTGCTCTATGACCTGGGACATCTGGAAGGTA	937
Db	241	tctcagctgcagcagagctgctcgtgctctatgacctggacatctlggaaggtta	300
Qy	938	CCCCATGGCCTTAGGAACTGGCAGATCTAGAGAGCTGAGGCCACCCCTGGCCGCC	997
Db	301	ccccatggccttaggaacctggcagatctagaggagctggagcccccctggccgccc	360
Qy	998	AGACCCACTCACCTCTACCAAGGCAAGGCTTCCTTCAGCCAAAGACCTACTATCGGATGA	1057
Db	361	agacccactcacctctaccacaaggcattgcttcggccaaagacctactatcggatga	420
Qy	1058	ACACATACCCCTACATGTACCTGGCTGGCTACCACTGTGCGCA 1101	
Db	421	acacatcacccctacatgtacctggctggctacctggtgctacctgtgcga 464	

RESULT 13

PCT-US01-00663-5551/c
; Sequence 5551, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
; FILE REFERENCE: PB 0004 WO 7

; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1003)...(1386)
; OTHER INFORMATION: similar to g13002527 in the genepept database release 115,
; OTHER INFORMATION: Run with FASTX 3.3t00, default parameters
US-09-577-408-6141

Query Match 16.18; Score 446; DB 22; Length 2121;
Best Local Similarity 100.0%; Pred. No. 2.7e-81;
Matches 446; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	111	ATGGGGCTGAAGCGCGCCAGAACACGCTGTTCCCGCTGGCGCTCCATCGACGACGTGGTG	170
Db	2121	ATGGGGCTGAAGCGCGCCAGAACACGCTGTTCCCGCTGGCGCTCCATCGACGACGTGGTG	2062
Qy	171	CGCTGTGTTGCTGCCGAGCTGGGCGGAGAGAGCGGACCTGGTGCTCCTTTCCCTGGTG	230
Db	2061	CGCTGTGTTGCTGCCGAGCTGGGCGGAGAGAGCGGACCTGGTGCTCCTTTCCCTGGTG	2002
Qy	231	CTGGGCTTCGTGGAGCATTTCTGGCTGTCAACCGCTCATCCCTACCAAGGTTCCCGAG	290
Db	2001	CTGGGCTTCGTGGAGCATTTCTGGCTGTCAACCGCTCATCCCTACCAAGGTTCCCGAG	1942
Qy	291	CTCACCTTCCAGCCAGCCCGCCCGAGCCCGCTGGCGGCTCAGCTACTTTCCCGTG	350
Db	1941	CTCACCTTCCAGCCAGCCCGCCCGAGCCCGCTGGCGGCTCAGCTACTTTCCCGTG	1882
Qy	351	GCCGACCTGTCTATCATCGCGCGCCCTCTATGCCGCTTACCGCCAGATCCGAGCGCC	410
Db	1881	GCCGACCTGTCTATCATCGCGCGCCCTCTATGCCGCTTACCGCCAGATCCGAGCGCC	1822
Qy	411	GTCCACCTGTCCCTCTATCTCGAGAGGGGGTGTCTCCAGCCGTGAGTGGTGAAGAAG	470
Db	1821	GTCCACCTGTCCCTCTATCTCGAGAGGGGGTGTCTCCAGCCGTGAGTGGTGAAGAAG	1762
Qy	471	GTCCTCCGATGTCATATGGAACAGCGCTCAGCGCTCCTACTTCAAGGATCGGGCCACATC	530
Db	1761	GTCCTCCGATGTCATATGGAACAGCGCTCAGCGCTCCTACTTCAAGGATCGGGCCACATC	1702
Qy	531	CAGTCCCTCTTCAGCTTCATCACAGG	556
Db	1701	CAGTCCCTCTTCAGCTTCATCACAGG	1676

Search completed: October 14, 2001, 07:39:50
Job time: 37716 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 20:00:54 ; Search time 218.79 Seconds
(without alignments)
2398.512 Million cell updates/sec

Title: US-09-380-337-1

Perfect score: 2772

Sequence: 1 GGTGTCGGAGCCGCGACC.....TGTTAAAAA.....2772

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA.*

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*

2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*

3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*

4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*

5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq.*

6: /cgn2_6/ptodata/1/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1865.2	67.3	2196	2	US-08-865-337A-2
2	863.8	31.2	986	2	US-08-865-337A-4
3	746.4	26.9	812	2	US-08-865-337A-5
4	492.2	17.8	538	2	US-08-865-337A-3
5	62.4	2.3	7218	1	US-08-232-463-14
6	57.2	2.1	1931	2	US-09-130-114-2
7	55.4	2.0	7218	1	US-08-232-463-14
8	52.6	1.9	38506	3	US-09-320-878-19
9	47.6	1.7	30001	1	US-08-125-468-1
10	47.6	1.7	30001	2	US-08-474-933-1
11	45	1.6	2370	1	US-08-104-072B-7
12	45	1.6	2370	1	US-08-351-413-8
13	45	1.6	2370	2	US-09-025-583-8
14	44.8	1.6	2219	3	US-08-510-646B-17
15	44.2	1.6	1146	1	US-08-482-385A-1
16	44.2	1.6	1780	2	US-08-933-821-5
17	44.2	1.6	1780	3	US-08-960-507-5
18	44.2	1.6	2728	1	US-08-482-385A-5
19	44.2	1.6	4524	2	US-08-845-998-7
20	44.2	1.6	4524	3	US-09-206-537-7
21	43.6	1.6	2580	3	US-09-050-863-2
22	43.6	1.6	2582	2	US-08-480-994-6
23	43.6	1.6	2582	2	US-08-616-844-6
24	43.6	1.6	2582	2	US-08-599-654-6
25	43.6	1.6	2582	2	US-08-485-573-6
26	43.6	1.6	2582	3	US-08-944-868A-6
27	43.6	1.6	2582	3	US-08-944-423A-6

28	43.6	1.6	2582	3	US-08-925-743-6	Sequence 6, Appl
29	43.6	1.6	2582	3	US-08-944-496-6	Sequence 6, Appl
30	43.6	1.6	2582	4	US-08-925-767-6	Sequence 6, Appl
31	43.6	1.6	5452	2	US-09-130-114-1	Sequence 1, Appl
32	43.6	1.6	9600	4	US-08-910-647-1	Sequence 1, Appl
33	43.6	1.6	10596	1	US-07-884-811-15	Sequence 15, Appl
34	43.6	1.6	10596	1	US-07-885-971-15	Sequence 15, Appl
35	43.6	1.6	10596	1	US-08-087-783A-15	Sequence 15, Appl
36	43.6	1.6	10596	1	US-08-194-088B-15	Sequence 15, Appl
37	43.6	1.6	10596	2	US-08-194-087-15	Sequence 15, Appl
38	43.6	1.6	10596	5	PCT-US93-04648-15	Sequence 15, Appl
39	43.2	1.6	1722	4	US-09-385-028-15	Sequence 15, Appl
40	43.2	1.6	11604	4	US-09-385-028-13	Sequence 13, Appl
41	43.2	1.6	15079	4	US-09-385-028-1	Sequence 1, Appl
42	43	1.6	2712	3	US-09-025-691-4	Sequence 4, Appl
43	43	1.6	12001	1	US-08-458-568A-11	Sequence 11, Appl
44	42	1.5	409	1	US-07-626-618A-8	Sequence 8, Appl
45	42	1.5	409	1	US-08-333-977-8	Sequence 8, Appl

ALIGNMENTS

RESULT 1
US-08-865-337A-2
; Sequence 2, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Kovitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2196 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT14
; CLONE: Consensus
; US-08-865-337A-2

Query Match 67.3%; Score 1865.2; DB 2: Length 2196; Best Local Similarity 91.4%; Pred. No. 0; Matches 2082; Conservative 7; Mismatches 50; Indels 140; Gaps 6;									
QY	88	GC	GC	GC	GC	GC	GC	GC	GC
DB	35	GC	GC	GC	GC	GC	GC	GC	GC
QY	148	TG	GC	GC	GC	GC	GC	GC	GC
DB	95	TG	GC	GC	GC	GC	GC	GC	GC
QY	208	AC	TC	GC	GC	GC	GC	GC	GC
DB	155	AC	TC	GC	GC	GC	GC	GC	GC
QY	268	TC	AT	CC	TC	AC	AG	TC	GC
DB	215	TC	AT	CC	TC	AC	AG	TC	GC
QY	328	GC	GC	GC	GC	GC	GC	GC	GC
DB	275	GC	GC	GC	GC	GC	GC	GC	GC
QY	388	TC	AC	GC	GC	GC	GC	GC	GC
DB	335	TC	AC	GC	GC	GC	GC	GC	GC
QY	448	CC	AG	CC	TC	AG	TC	GC	GC
DB	395	CC	AG	CC	TC	AG	TC	GC	GC
QY	508	AC	TC	AA	GC	GC	GC	GC	GC
DB	455	AC	TC	AA	GC	GC	GC	GC	GC
QY	568	AC	AG	TC	GC	GC	GC	GC	GC
DB	515	A	---	CT	CG	GC	GC	GC	GC
QY	628	TC	AC	TC	GC	GC	GC	GC	GC
DB	535	---	---	---	---	---	---	---	---
QY	688	C	AG	TC	GC	GC	GC	GC	GC
DB	560	C	AG	TC	GC	GC	GC	GC	GC
QY	748	CC	GC	TC	GC	GC	GC	GC	GC
DB	620	CC	GC	TC	GC	GC	GC	GC	GC
QY	808	AG	TC	GC	GC	GC	GC	GC	GC
DB	680	AG	TC	GC	GC	GC	GC	GC	GC
QY	868	CG	TC	GC	GC	GC	GC	GC	GC
DB	740	CG	TC	GC	GC	GC	GC	GC	GC
QY	928	TG	AA	AG	TC	GC	GC	GC	GC
DB	800	TG	AA	AG	TC	GC	GC	GC	GC
QY	988	CT	GC	GC	GC	GC	GC	GC	GC
DB	860	CT	GC	GC	GC	GC	GC	GC	GC
QY	1048	AT	CG	GC	GC	GC	GC	GC	GC
DB	920	AT	CG	GC	GC	GC	GC	GC	GC

QY	1108	AT	TC	GC	GC	GC	GC	GC	GC
DB	980	AT	TC	GC	GC	GC	GC	GC	GC
QY	1165	AC	TC	GC	GC	GC	GC	GC	GC
DB	1040	AC	TC	GC	GC	GC	GC	GC	GC
QY	1225	TC	CC	AA	CC	TC	GC	GC	GC
DB	1100	TC	CC	AA	CC	TC	GC	GC	GC
QY	1285	AG	CA	AA	CC	TC	GC	GC	GC
DB	1160	AG	CA	AA	CC	TC	GC	GC	GC
QY	1345	AC	TC	GC	GC	GC	GC	GC	GC
DB	1220	AC	TC	GC	GC	GC	GC	GC	GC
QY	1405	TG	CA	GC	TC	GC	GC	GC	GC
DB	1280	TG	CA	GC	TC	GC	GC	GC	GC
QY	1465	GC	CA	AA	AG	TC	GC	GC	GC
DB	1340	GC	CA	AA	AG	TC	GC	GC	GC
QY	1525	GC	CA	AA	AG	TC	GC	GC	GC
DB	1381	---	---	---	---	---	---	---	---
QY	1585	CC	GC	GC	GC	GC	GC	GC	GC
DB	1400	CC	GC	GC	GC	GC	GC	GC	GC
QY	1645	C	AG	TC	GC	GC	GC	GC	GC
DB	1460	C	AG	TC	GC	GC	GC	GC	GC
QY	1705	GC	CA	GC	TC	GC	GC	GC	GC
DB	1520	GC	CA	GC	TC	GC	GC	GC	GC
QY	1765	CT	TC	CA	GC	TC	GC	GC	GC
DB	1580	CT	TC	CA	GC	TC	GC	GC	GC
QY	1825	CG	AG	CC	TC	GC	GC	GC	GC
DB	1640	CG	AG	CC	TC	GC	GC	GC	GC
QY	1885	TG	TC	CA	GC	TC	GC	GC	GC
DB	1700	TG	TC	CA	GC	TC	GC	GC	GC
QY	1945	CT	AG	TC	GC	GC	GC	GC	GC
DB	1760	CT	AG	TC	GC	GC	GC	GC	GC
QY	2004	G	AG	CC	TC	GC	GC	GC	GC
DB	1820	G	AG	CC	TC	GC	GC	GC	GC
QY	2064	TT	CC	GC	TC	GC	GC	GC	GC
DB	1880	TT	CC	GC	TC	GC	GC	GC	GC
QY	2124	G	CT	TC	GC	GC	GC	GC	GC
DB	1940	G	CT	TC	GC	GC	GC	GC	GC
QY	2184	A	AC	TC	GC	GC	GC	GC	GC

Db 2000 AACCTCACGCTGGTCTAATTACCAAGTCTTTAAAGGCCAGCCCTAGAAACCAAGCT 2059
QY 2244 CTCTCTCGG-AACCGCTACCTAGAGCCAGACCAAGCTTACTCAGGCTCTCCAGCTT 2302
Db 2060 CNTCTCGGNAACCGTACCTAGAGCCAGACCAAGCTTAYTCAGGGGTCCTNCCAGCTT 2119
QY 2303 GTAGGAGCTGAGGTTTACCCCTTAACCCAAAGGAGCAGACAGGTCCTCCACCTCCAGCCCGGG 2361
Db 2120 GTAGGAGCTGAGGTTTACCCCTTAACCCAAAGGAGGAAAGGTTCCCANCTTCAGCCCGGG 2178

RESULT 2

US-08-865-337A-4
; Sequence 4, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 986 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BLADTUT08
; CLONE: 2809030
US-08-865-337A-4

Query Match 31.2%; Score 863.8; DB 2; Length 986;
Best Local Similarity 92.9%; Pred. No. 1.9e-198;
Matches 907; Conservative 2; Mismatches 63; Indels 4; Gaps 2;
QY 668 TGGGCCCAATGGGGAGCAGACAGCTGAGTCCACCTGGCAGCGGCAAGGCAAGCAGGACCG 727
Db 1 TGGGCCCAATGGGGAGCAGACAGCTGAGTCCACCTGGCAGCGGCAAGGCAAGCAGGACCG 60
QY 728 CAGGCGCCAGACAGTCAATGCCGGTGTGGCTGAGCGGAGCTGGCTACTGAAAGATC 787

Db 61 GAGGGCCAGACAGTCAATGCCGGTGTGGCTGAGCGGAGCTGGCTGTACC:GAAGGATC 120
QY 788 ATACATGCGCTGTGACCGCAAGATGGAGTGGGCTTCATCGTGTGTCCTCAACCCCTTC 847
Db 121 ATACATGCGCTGTGACCGCAAGATGGAGTGGGCTTCATCGTGTGTCCTCAACCCCTTC 180
QY 848 CATTTGACCTGTCACACCGGACTTCGCTGGAGCTTCTGCAGCTGCAGCAGAAGTGTCTCTGGCT 907
Db 181 CATTTGACCTGTCACACCGGACTTCGCTGGAGCTTCTGCAGCTGCAGCAGAAGTGTCTCTGGCT 240
QY 908 GCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCCCTTAGGGAACCTGGCAGATCT 967
Db 241 GCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCCCTTAGGGAACCTGGCAGATCT 300
QY 968 AGAGGAGCTGGAGCCACCCCTGCGCGGCCAGACCCACTCACCTCTTACACAGGCGCAT 1027
Db 301 AGAGGAGCTGGAGCCACCCCTGCGCGGCCAGACCCCTGCGCGGCCCTGCTRCACAGGCGCAT 360
QY 1028 TGCTTCAGCCCAAGACCTTACTATCGGGATGAACACATCTACCCCTTACATGTACCTGGCTGG 1087
Db 361 TGCTTCAGCCCAAGACCTTACTATCGGGATGAACACATCTACCCCTTACATGTACCTGGCTGG 420
QY 1088 CTACCACTGTGCAACCGCAATGTGCGGGAAGCCCTGCGAGGCTGGG---CGGACACGGC 1144
Db 421 CTACCACTGTGCAACCGCAATGTGCGGGAAGCCCTGCGAGGCTGGG---CGGACACGGC 480
QY 1145 CACTGTCTACCCAGGACTACCACTACTGCGCGGAGAGGAGGAGATCTACAGGAGTCTTT 1204
Db 481 CACTGTCTACCCAGGACTACCACTACTGCGCGGAGAGGAGGAGATCTACAGGAGTCTTT 540
QY 1205 TGAAGTAGCAATGATGTCATCCCAACCTGCTGAAGGAGCGGAGCCAGCTTGTGGAGGC 1264
Db 541 TGAAGTAGCAATGATGTCATCCCAACCTGCTGAAGGAGCGGAGCCAGCTTGTGGAGGC 600
QY 1265 GCGGAGGAGCGCGCGGGGAGCAAGCCAGGGGACCCAGAGCCCAAGGTTCCGCGCTCCA 1324
Db 601 GCGGAGGAGCGCGCGGGGAGCAAGCCAGGGGACCCAGAGCCCAAGGTTCCGCGCTCCA 660
QY 1325 GGACCTGTAGTGTCTGCGCCACCTGCTGCATTTCTACGACGGCATCTGCAAAATGGGAGGA 1384
Db 661 GGACCTGTAGTGTCTGCGCCACCTGCTGCATTTCTACGACGGCATCTGCAAAATGGGAGGA 720
QY 1385 GGCAGTCCCAAGCCCTGTGTCACGCTGGGCTGGGCGCACCTTTCTTGTGTCAGTCCCTTAGG 1444
Db 721 GGCAGTCCCAAGCCCTGTGTCACGCTGGGCTGGGCGCACCTTTCTTGTGTCAGTCCCTTAGG 780
QY 1445 CGTTTGTAGGGACAGGTGCGGCGAGAGGTGCGCATAGTAGCGGAGAGCGGCGGCGGCG 1504
Db 781 CGTTTGTAGGGACAGGTGCGGCGAGAGGTGCGCATAGTAGCGGAGATGCCGAGGCTGC 840
QY 1505 CGAGGCGGAGGAGCGGTGGGGCGAGGAAGCCCGGAGGCGCGGCGGCGGCGGCGGCGGCGG 1564
Db 841 CGAGGCGG-ANGAGCGGTGGNGCTAGGAAGCCCGGNAAGCCCGGNGNGGCGGCGGCGGCGG 899
QY 1565 GGAGTCCCAAGCCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 1624
Db 900 GGAGTCCCAAGCCAGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 959
QY 1625 CACCGCGGAGGTTGCA 1640
Db 960 CACCGCGGAGGTTGCA 975

RESULT 3
US-08-865-337A-5
; Sequence 5, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE

;; TITLE OF INVENTION: 1 PROTEIN
;; NUMBER OF SEQUENCES: 6
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Incyte Pharmaceuticals, Inc.
;; STREET: 3174 Porter Drive
;; CITY: Palo Alto
;; STATE: CA
;; COUNTRY: USA
;; ZIP: 94304
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Diskette
;; OPERATING SYSTEM: DOS
;; SOFTWARE: FastSeq for Windows Version 2.0
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/865,337A
;; FILING DATE: Herewith
;; CLASSIFICATION: 530
;; PRIOR APPLICATION NUMBER:
;; APPLICATION NUMBER:
;; FILING DATE:
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Billings, Lucy J.
;; REGISTRATION NUMBER: 36,749
;; REFERENCE/DOCKET NUMBER: PF-0305 US
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 415-855-0555
;; TELEFAX: 415-845-4166
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 5:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 812 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; IMMEDIATE SOURCE:
;; LIBRARY: BLADTUT02
;; CLONE: 1313372
US-08-865-337A-5

Query Match 26.9%; Score 746.4; DB 2; Length 812;
Best Local Similarity 97.2%; Pred. No. 2.7e-170;
Matches 772; Conservative 4; Mismatches 16; Indels 2; Gaps 2;
QY 1570 CCAAGCCAGAGGAGCCCGCCGCCCAAGAGCCAGCAGCTGGACAGGGGCTGGGACCG 1629
DB 1 CCAAGCCAGAGGAGCCCGCCGCCCAAGAGCCAGCAGCTGGACAGGGGCTGGGACCG 60
QY 1630 GCCAGGCTGAGTGTCTAGGACCCCGCCGGAAGCTCTCTGGGACTGTCTGGGACAGCC 1689
DB 61 GCCAGGCGCAGTGTCTAGGACCCCGCCGGAAGCTCTCTGGGACTGTCTGGGACAGCC 120
QY 1690 GAGGCCCTGAGGTGGCAGCAGCGCTCAGTGGCAGCAGCCCGCAGCATCACCGCCCG 1749
DB 121 GAGGCCCTGAGGTGGCAGCAGCGCTCAGTGGCAGCAGCCCGCAGCATCACCGCCCG 180
QY 1750 AGGTCCAGTGTCTCAGTGTCTCAGAGTGAAGATGAAGGGCATGAAGGAGCTGTCTGGTGG 1809
DB 181 AGGTCCAGTGTCTCAGTGTCTCAGAGTGAAGATGAAGGGCATGAAGGAGCTGTCTGGTGG 240
QY 1810 CCACCAAGATCAACTCGAGGCGCCATCAAGTGAACCTCAGCGGCACATCGCAAGTGCAGA 1869
DB 241 CCACCAAGATCAACTCGAGGCGCCATCAAGTGAACCTCAGCGGCACATCGCAAGTGCAGA 300
QY 1870 TGAGAGAGCAGAAAGTGTCCACCCCTAGTGTCTCTCTTCTCTCAAGGGGCGAG 1929
DB 301 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGTCTCTCTTCTCTCAAGGGGCGAG 360
QY 1930 GCAAGGCCCTCTGAACCTACTGTGGGACTTCGGACCGCTGTGGGACCCAGCGCTCCG-CCT 1988
DB 361 CAAAGGCCCTCTGAACCTACTGTGGGACTTCGGACCGCTGTGGGACCCAGCGCTCCGCGCT 420

QY 1989 TAGTCCCCCAACTCTCAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGGCTCACCTC 2048
DB 421 TAGTCCCCCAACTCTCAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGGCTCACCTC 480
QY 2049 TACCCAAACCCCTAGTGTCCCGGTCGCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 2108
DB 481 TACCCAAACCCCTAGTGTCCCGGTCGCCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCC 540
QY 2109 AGCTCAGAACCCAGGCGCTCTGCCCCAGTGGTGTAGATATAGGTCTCTCTCTCCAGATCC 2168
DB 541 AGCTCAGAACCCAGGCGCTCTGCCCCAGTGGTGTAGATATAGGTCTCTCTCTCCAGATCC 600
QY 2169 CAGCCGGCCCAATGGAACCTCAGCTGGTGGTCTTAATACAGTCTTTAAAGGCCAGCC 2228
DB 601 CAGCCGGCCCAATGGAACCTCAGCTGGTGGTCTTAATACAGTCTTTAAAGGCCAGCC 660
QY 2229 CTAGAAACCCAGCTCTCTCTCGG-AACCGCTCAGCTAGAGCCAGACCAACGTTACTCAG 2287
DB 661 CTAGAAACCCAGCTCTCTCTCGG-AACCGCTCAGCTAGAGCCAGACCAACGTTACTCAG 720
QY 2288 GGCTCTCTCCAGCTTGTAGGAGCTGAGGTTTACCTTTAACCCAGGGAGCAGGTTCC 2347
DB 721 GGCTCTCTCCAGCTTGTAGGAGCTGAGGTTTACCTTTAACCCAGGGAGGAAAGGTCC 780
QY 2348 ACCTCCAGCCCGG 2361
DB 781 ANCTTCAGCCCGG 794
RESULT 4
US-08-865-337A-3
; Sequence 3, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single

```
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT14
; CLONE: 1596949
US-08-863-337A-3

Query Match      17.8%; Score 492.2; DB 2; Length 538;
Best Local Similarity 99.2%; Pred. No. 2.6e-109;
Matches 502; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 88 GCGCGCGCCACCGCGCGCGCATGGGCTGAAGCGCCGAGAGAGCGCTGTTCGCCG 147
Db 34 GCGCGCGCCACCGCGCGCGCATGGGCTGAAGCGCCGAGAGAGCGCTGTTCGCCG 93

QY 148 TCGCGTCCATCAGCAGCGTGGCGCTGCTTGGCTGCGGAGCTGGCGCCGAGAGGCGCG 207
Db 94 TCGCGTCCATCAGCAGCGTGGCGCTGCTTGGCTGCGGAGCTGGCGCCGAGAGGCGCG 153

QY 208 ACCTGGTGTCTCTTCCCTTGGTGGCTTGGTGGAGCATTTTCTGGCTGTCAACCGCG 267
Db 154 ACCTGGTGTCTCTTCCCTTGGTGGCTTGGTGGAGCATTTTCTGGCTGTCAACCGCG 213

QY 268 TCATCCCTACCAACGTTCCCGAGCTACCTTCCAGCCCGCGCCGCGCCGCGCGCTG 327
Db 214 TCATCCCTACCAACGTTCCCGAGCTACCTTCCAGCCCGCGCGCCGCGCGCGCTG 273

QY 328 GCGGCTACCTTCTCCGTTGGCGGAGCTGTCTATCATCGCGCGCTCTATGCGCGCT 387
Db 274 GCGGCTACCTTCTCCGTTGGCGGAGCTGTCTATCATCGCGCGCTCTATGCGCGCT 333

QY 388 TCACGCGCCAGATCCGAGCGCGCTGAGCTGTCTCTCTATCTCGAGAGGGGTGTCT 447
Db 334 TCACGCGCCAGATCCGAGCGCGCTGAGCTGTCTCTCTATCTCGAGAGGGGTGTCT 393

QY 448 CCAGCGGTGAGTGTGAAGAAGGTCTCGATGTGCATATGGAACACGCTTCAGCGCTCT 507
Db 394 CCAGCGGTGAGTGTGAAGAAGGTCTCGATGTGCATATGGAACACGCTTCAGCGCTCT 453

QY 508 ACTTCAAGATCGGCGCCACATCCAGTCCCTCTTCAGCTTCATCAGGACCAATTCG 567
Db 454 ACTTCAAGATTCGGCGCCACATCCAGTCCCTCTTCAGCTTCATCAGGACCAATTCG 513

QY 568 ACAGTCCGCTGGCGCTTGTGTGTG 593
Db 514 ACA-STCCGNGTGGCGCTTGTGTGTG 538

RESULT 5
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
```

```
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match      2.3%; Score 62.4; DB 1; Length 7218;
Best Local Similarity 6.4%; Pred. No. 1.4e-05;
Matches 24; Conservative 207; Mismatches 143; Indels 0; Gaps 0;

QY 178 TTGCTGCCGAGCTGGCGCGAGAGCGCCGACCTGTGCTGCTCTCTTCTGCTGCGGT 237
Db 1027 TTAATTCCGAGCTTGCTCGAGGTGAGGAGCTTGGATTTTTTTTTTTTTTTTTT 1086

QY 238 TCGTGAGGAGATTTTGGCTGTCAACCGGTGTATCCCTACCAAGTTCCGAGCTCACCT 237
Db 1087 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1146

QY 298 TCACGCGCGAGCGCGCCCGCGCCGCTGGCGGCTCACCTACTTCCGTCGCGGACC 357
Db 1147 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1206

QY 358 TGCTATATATCGCGCGCTCTATGCGCGCTTACCGCGGAGATCCGAGCGCGCTGAC 417
Db 1207 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1266

QY 418 TGTCCCTCTATCTCGAGAGGGGTGTCTCCAGCGGTGAGCTGGTGAAGAAGTCTCCG 477
Db 1267 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1326

QY 478 ATGTCATATGGAACAGCTCAGCGCTCTCTACTTCAAGGATCGGCGCCACATCCAGTCC 537
Db 1327 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1386

QY 538 TCTTCAGTTCATC 551
Db 1387 YYYYYYYYYYYYYY 1400

RESULT 6
US-09-130-114-2
; Sequence 2, Application US/09130114
; Patent No. 5976807
; GENERAL INFORMATION:
; APPLICANT: Horlick, Robert A.
; APPLICANT: Dama, Bassam B.
; APPLICANT: Robbins, Alan K.
; TITLE OF INVENTION: Eukaryotic Cells Stably Expressing Genes
; TITLE OF INVENTION: From Multiple Transfected Episomes
; FILE REFERENCE: 0867/1D903US1
; CURRENT APPLICATION NUMBER: US/09/130,114
; CURRENT FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 36
; SOFTWARE: FastSeq for Windows Version 3.0
```


RESULT 10


```
Db 1875 GCGGGGGCGCGTGGCGGGGTGTCAGTGGCGGCGAGGTGATGAGCTGATGGCGCGCTGC 1934
QY 225 TTGGTCTGGGCTTCGTGGAGCATTTTCTGGGTGTCAACCGGGGTATCCCTACCAACGTT 284
Db 1935 ATGCGGTACCTCGCGGCGCGCGGGATGACGCCCTACGGCATCTGCTGGACAGCCTC 1994
QY 285 CCGGAGCTCACCTTCAGCCAGCCCGCGCGGAGCCCGCTGGGGGCTCACCTACTTTT 344
Db 1995 GCGGTCTCAACGGATGTCGCGCGCGCGCGGCGGCGGCGGCTCTGCAACTGCGTC 2054
QY 345 CCGGTGGCGGACCTGTCTATCATCGCGCGCTCTATGATCGCGGCTTACCGCCAGATCCGA 404
Db 2055 AAGGAGCGCGCGCGGCTTCCCGCGGTGACCTTCCCGCGCTCCGCC---TCCCC 2111
QY 405 GCGCGCTGCGACCTGTCCCTC 425
Db 2112 GCGCGCTGCGCTCTCCATC 2132

RESULT 12
US-08-351-413-8
: Sequence 8, Application US/08351413
: Patent No. 5750867
: GENERAL INFORMATION:
: APPLICANT: Williams, Mark
: APPLICANT: Leemans, Jan
: TITLE OF INVENTION: Maintenance of male-sterile plants
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 2046
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: FILING DATE:
: APPLICATION NUMBER: US/08/351,413
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/899,072
: FILING DATE: 12-JUN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/970,849
: FILING DATE: 03-NOV-1992
: ATTORNEY/AGENT INFORMATION:
: NAME: Svensson, Leonard R.
: REGISTRATION NUMBER: 30,330
: REFERENCE/DOCKET NUMBER: 2121-102PCT
: TELEPHONE: (703) 205-8000
: TELEFAX: (703) 205-8050
: TELEX: 248345
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2370 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: HYPOTHETICAL: NO
: ANTI-SENSE: NO
: ORIGINAL SOURCE:
: ORGANISM: Oryza sativa
: STRAIN: Akihikari
: FEATURE:
: NAME/KEY:
: LOCATION: 1..1808

Db 1875 GCGGGGGCGCGTGGCGGGGTGTCAGTGGCGGCGAGGTGATGAGCTGATGGCGCGCTGC 1934
QY 225 TTGGTCTGGGCTTCGTGGAGCATTTTCTGGGTGTCAACCGGGGTATCCCTACCAACGTT 284
Db 1935 ATGCGGTACCTCGCGGCGCGCGGGATGACGCCCTACGGCATCTGCTGGACAGCCTC 1994
QY 285 CCGGAGCTCACCTTCAGCCAGCCCGCGCGGAGCCCGCTGGGGGCTCACCTACTTTT 344
Db 1995 GCGGTCTCAACGGATGTCGCGCGCGCGCGGCGGCGGCTCTGCAACTGCGTC 2054
QY 345 CCGGTGGCGGACCTGTCTATCATCGCGCGCTCTATGATCGCGGCTTACCGCCAGATCCGA 404
Db 2055 AAGGAGCGCGCGCGGCTTCCCGCGGTGACCTTCCCGCGCTCCGCC---TCCCC 2111
QY 405 GCGCGCTGCGACCTGTCCCTC 425
Db 2112 GCGCGCTGCGCTCTCCATC 2132

Query Match 1.6%; Score 45; DB 1; Length 2370;
Best Local Similarity 51.0%; Pred. No. 0.13;
Matches 133; Conservative 0; Mismatches 125; Indels 3; Gaps 1;
QY 165 GTGGTGGCGCTGTTCGTGGCGGAGTGGCGGAGAGGAGCGCGACCTGTGCTCTCTTTCC 224
Db 1875 GCGGGGGCGCGTGGCGGGGTGTCAGTGGCGGCGAGGTGATGAGCTGATGGCGCGCTGC 1934
QY 225 TTGGTCTGGGCTTCGTGGAGCATTTTCTGGGTGTCAACCGGGGTATCCCTACCAACGTT 284
Db 1935 ATGCGGTACCTCGCGGCGCGCGGGATGACGCCCTACGGCATCTGCTGGACAGCCTC 1994
QY 285 CCGGAGCTCACCTTCAGCCAGCCCGCGCGGAGCCCGCTGGGGGCTCACCTACTTTT 344
Db 1995 GCGGTCTCAACGGATGTCGCGCGCGCGCGGCGGCTCTGCAACTGCGTC 2054
QY 345 CCGGTGGCGGACCTGTCTATCATCGCGCGCTCTATGATCGCGGCTTACCGCCAGATCCGA 404
Db 2055 AAGGAGCGCGCGCGGCTTCCCGCGGTGACCTTCCCGCGCTCCGCC---TCCCC 2111
QY 405 GCGCGCTGCGACCTGTCCCTC 425
Db 2112 GCGCGCTGCGCTCTCCATC 2132

RESULT 13
US-09-025-583-8
: Sequence 8, Application US/09025583
: Patent No. 5977433
: GENERAL INFORMATION:
: APPLICANT: Williams, Mark
: APPLICANT: Leemans, Jan
: TITLE OF INVENTION: Maintenance of male-sterile plants;
: NUMBER OF SEQUENCES: 17
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: BIRCH, STEWART, KOLASCH & BIRCH
: STREET: 8110 Gatehouse Road, Suite 500 East
: CITY: Falls Church
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 2046
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25 (EPO)
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/025,583
: FILING DATE:
: CLASSIFICATION: 800
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/351,413
```

FILING DATE: US 07/899,072
FILING DATE: 12-JUN-1992
PRIOR APPLICATION DATA: US 07/970,849
FILING DATE: 03-NOV-1992
ATTORNEY/AGENT INFORMATION:
NAME: Svensson, Leonard R.
REGISTRATION NUMBER: 30,330
REFERENCE/DOCKET NUMBER: 2121-102PCT
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX: 248345
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 2370 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Oryza sativa
STRAIN: Akihikari
FEATURE:
NAME/KEY: -
LOCATION: 1..1808
OTHER INFORMATION: /label= PT42
OTHER INFORMATION: /note= "sequence comprising anther specific promoter PT42"
FEATURE:
NAME/KEY: -
LOCATION: 1748..1755
OTHER INFORMATION: /label= TATA
OTHER INFORMATION: /note= "TATA Box"
FEATURE:
NAME/KEY: -
LOCATION: 1780
OTHER INFORMATION: /note= "transcription initiation site determined by primer extension"
FEATURE:
NAME/KEY: -
LOCATION: 1809
OTHER INFORMATION: /label= ATG
OTHER INFORMATION: /note= "ATG start of translation of rice T42 gene"
US-09-025-583-8

Query Match 1.6%; Score 45; DB 2; Length 2370;
Best Local Similarity 51.0%; Pred. No. 0.13;
Matches 133; Conservative 0; Mismatches 125; Indels 3; Gaps 1;
Qy 165 GTGGTGGCGCTGTTGTGCGGAGCTGGCGGAGAGGAGCGGACCTGGTGGCTCTTTC 224
Db 1875 GCGGGGGCGCGTGGCGGGGTTGTCAGTGGCGGAGGTGATGACGTGATGGCGCGCTG 1934
Qy 225 TTGGTGTGGCTGCTGGAGCATTTTCTGGCTGTCAACCGCGTCACTCCCTACCAAGTT 284
Db 1935 ATGCGGTACCTCGCGCGCGCGCGCGGATGACGCGCTACGCGCATCTGCTGCGACAGCTC 1994
Qy 285 CCCGAGCTACCTTCCAGCCAGCCCGCGCGCGCGCGCGCTGCGCGCGCTCACCTACTTT 344
Db 1995 GCGGTGTCAACCGGATGGCGCGCGCGCGCGCGCGCGCTGCGCGCTGCACTGGCTC 2054
Qy 345 CCGGTGGCGGACCTGTCTATCATCGCGCGCTCTATCGCGCTTACCGCGCGCGAGATCGGA 404
Db 2055 AAGGACGCGCGCGCGGCTTCCCGCGCGCTGCACTTCTCCGCGCGCTCCGCC---TCC 2111
Qy 405 GCGCGCGCTGACCTGCTCCCTC 425
Db 2112 GCCGCGTGGCGCGCTCTCCATC 2132

RESULT 14
US-08-510-646B-17
Sequence 17, Application US/08510646B
Patent No. 6077699
GENERAL INFORMATION:
APPLICANT: Blanc, Veronique
APPLICANT: Blanche, Francis
APPLICANT: Crouzet, Joel
APPLICANT: Jacques, Nathalie
APPLICANT: Lacroix, Patricia
APPLICANT: Thibaut, Denis
APPLICANT: Zagrec, Monique
APPLICANT: Debussche, Laurent
APPLICANT: De Crecy-Lagard, Valerie
TITLE OF INVENTION: Polypeptides Involved In The
Biosynthesis Of Streptogramins, Nucleotide Sequences
Coding For These Polypeptides And Their Use
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Finnegan, Henderson, Farabow, Garrett & Dunner
STREET: 1300 I Street, N.W., Suite 700
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20005-3315
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/510,646B
FILING DATE: 03-AUG-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/403,852
FILING DATE: 10-MAY-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/FR 93/00923
FILING DATE: 25-SEP-1993
PRIOR APPLICATION DATA:
APPLICATION NUMBER: FR 92/11441
FILING DATE: 25-SEP-1992
ATTORNEY/AGENT INFORMATION:
NAME: Meyers, Kenneth J.
REGISTRATION NUMBER: 25,146
REFERENCE/DOCKET NUMBER: 03806.0054-01000
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 408-4000
TELEFAX: (202) 408-4400
INFORMATION FOR SEQ ID NO: 17:
SEQUENCE CHARACTERISTICS:
LENGTH: 2219 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: S. virginiae
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2219
OTHER INFORMATION: /product = "virginiamycin s synthase gene"
US-08-510-646B-17

Query Match 1.6%; Score 44.8; DB 3; Length 2219;
Best Local Similarity 48.7%; Pred. No. 0.14;
Matches 186; Conservative 0; Mismatches 187; Indels 9; Caps 2;

	Best Local Similarity	43.8%;	Pred. No. 0.15;			
	Matches	242;	Conservative	Mismatches	308;	Indels
						3; Gaps
Qy	93	CGCCACCGCCGCGCCATGGGGCTGAAGCGCGCCCGACAGAGCGTGTTCGCGCTCGCG	152			
Db	151	CGCGCGTGTACGCGAGCTGGTGGCGCGCGCGTACAAACACGACGACGACGGCTCTC	210			
Qy	153	TCATCAGACAGCTGTGTGCGCTGTTTGTGTCCGAGCTGGCGCGAGGAGCGGACCTG	212			
Db	211	ACCAAGAGGGCGCGCTCGCGCTGTACCCGTTCGAGCACGGGCGC---TAGAGGCTCGAG	267			
Qy	213	GTGTCCTTTCTTGTGTGTGGGCTTCGTGGAGCATTTTCTGTCTGTCAACCGCGTTCATC	272			
Db	268	GTGCGCGCGGCTCGTGTCTGGAGGACGCGACTGGCTCTTCCCAGCTACCGGGACAC	327			
Qy	273	CTTACCAAGCTTTCGAGCTCACTTTCACGCGCAGCCCCCGCCCGCCGCTTGGGGCG	332			
Db	328	CTCGCCCGCTCGCGCGCGGCTCGAATCCGCTGCAGCGCTCACTCTCTTCGCGGGCGAC	387			
Qy	333	CTCACCTACTTTTCCGTGGCGGACCTCTCTATCATCGCCGCCCTCTATGCCCGCTTCACC	392			
Db	388	TGGCACACCGGGTACGACCCCGTGAGCACCGCATCGCGCCCTGTGCACCCCTCTCGG	447			
Qy	393	GCCAGATTCGAGGCGCGCTGCAGCTGTCCCTTATCTCTGAGAGGGGGTGTCTCCAGC	452			
Db	448	ACCAGCTCCCGCACGCGCTGGCGCTCGCGACGCGCGCCGCTCAAGGCGACGACGTG	507			
Qy	453	CGTCAGCTGTGAAGAGGTCTCCGATGTATATGGAACAGCTCAAGCCGCTCCTACTTC	512			
Db	508	GTGCGGTTCGCCCTGTCTGGGGACGGCGGACACGAGGAGGCGACATTCACGAGGACCTG	567			
Qy	513	AAGGATCGGGGCCACATCCACTCCCTCTCAGCTTCATCACAGGCACCAAATTTGGACAGC	572			
Db	568	AACCTTCGCCCGCGTCTGGCAGGGCGCGGTCTTCTCTCGTGTGAGAACAAACGGCTTCGCC	627			
Qy	573	TCGCGTGTGGCCTTTGTGTGTGGGCGTTCGAGCGCCTGTGGTCTTCGCGGATGTCCAC	632			
Db	628	ATCTCGTCCGCTCGCAAGCAGACACCGCGCCCGCTGCTGSCCCCAACAGGCGCTCGGC	687			
Qy	633	CTCGCCCTGTCTG	645			
Db	688	TACGGGATGCGG	700			

Query Match 1.68; Score 44.2; DB 1; Length 1146;

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 21:01:20 ; Search time 644.85 seconds
(without alignments)
2699.144 Million cell updates/sec

Title: US-09-380-337-1

Perfect score: 2772

Sequence: 1 GGTGTCGGAGCCGCGACC.....TGTTAAAAA..... 2772

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

```

1: /SIDS8/gcgdata/geneseq/geneseqn/NA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseqn/NA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseqn/NA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseqn/NA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseqn/NA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseqn/NA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseqn/NA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseqn/NA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseqn/NA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseqn/NA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseqn/NA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseqn/NA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseqn/NA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseqn/NA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseqn/NA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseqn/NA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseqn/NA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseqn/NA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseqn/NA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseqn/NA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseqn/NA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseqn/NA2001.DAT.*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2772	100.0	2772	19 AAV57415	Tumour suppressor
2	1865.2	67.3	2196	20 AAV80659	Human multiple end
3	1312.2	47.3	9180	19 AAV57416	Tumour suppressor
4	863.8	31.2	986	20 AAV80661	Human multiple end
5	746.4	26.9	812	20 AAV80662	Human multiple end
6	492.2	17.8	538	20 AAV80660	Human multiple end
7	234	8.4	253	20 AAV86611	EST clone BE3. Ho
8	66	2.4	67	16 AAT22851	Human gene signatu
9	52.6	1.9	13842	21 AAZ87297	S. venezuelae macr
10	52.6	1.9	36778	21 AAZ87318	S. venezuelae pik
11	52.6	1.9	37948	21 AAZ87285	S. venezuelae pik

12	52.6	1.9	38506	21	AAV575633	Nucleotide sequenc
13	52.6	1.9	38506	21	AAZ56001	Recombinant cosmid
14	52.2	1.9	10732	21	AAAL0594	Gene encoding a su
15	52.2	1.9	15872	18	AAT68715	Streptomyces venez
16	52.2	1.9	15872	21	AAZ87283	S. venezuelae vep
c 17	50.8	1.8	114955	20	AAZ53491	Human adenosine Al
18	49.6	1.8	1725	21	AAAL0503	Trehalose-releasin
19	49.6	1.8	2218	21	AAAL0516	Human cytoskeletal
20	48.8	1.8	7962	21	AAZ58980	Sorghum bicolor (L
21	48.2	1.7	1929	19	AAV57472	Human adenosine Al
22	48	1.7	114955	20	AAZ53491	Nucleotide sequenc
c 23	47.8	1.7	5857	21	AAZ58471	Total DNA sequence
c 24	47.6	1.7	30001	18	AAZ61016	S. aureofaciens DN
c 25	47.6	1.7	30001	20	AAZ05110	Gene encoding a su
c 26	47	1.7	10732	21	AAAL0594	HSV-2 strain SB5 C
c 27	47	1.7	21034	19	AAV62154	HSV-2 strain SB5 C
c 28	47	1.7	117213	19	AAV62176	Nucleotide sequenc
29	45.8	1.7	18660	21	AAZ58472	Human type VI aden
30	45.4	1.6	4942	20	AAZ00461	Human G protein-co
31	45.2	1.6	1128	21	AAZ00619	DNA encoding human
32	45.2	1.6	1128	21	AAZ00724	Human CEPR (hCEPR)
33	45.2	1.6	1648	21	AAZ11700	PEG-1 gene. Homo
34	45.2	1.6	2776	18	AAZ73366	Sequence comprisin
35	45	1.6	2370	15	AAQ53880	S. clavuligerus cl
36	44.8	1.6	15079	16	AAQ91580	Streptomyces prist
37	44.2	1.6	1065	17	AAZ58554	S. avermitilis BCK
38	44.2	1.6	1146	16	AAZ03793	Human NEW angiopo
39	44.2	1.6	1413	22	AAZ26792	Human TIE ligand N
40	44.2	1.6	1780	20	AAZ36343	Human PRO178 CDNA
41	44.2	1.6	1780	21	AAZ77516	Human anglogenesis
42	44.2	1.6	1780	22	AAZ97375	S. avermitilis bkd
43	44.2	1.6	2728	16	AAZ03797	Streptomyces prist
44	44.2	1.6	4496	17	AAZ58553	Nucleotide sequenc
c 45	44.2	1.6	4524	20	AAV33912	

ALIGNMENTS

```

RESULT 1
AAV57415
ID AAV57415 standard; cDNA; 2772 BP.
AC
XX AAV57415;
XX
XX
DT 11-JAN-1999 (first entry)
XX
DE Tumour suppressor gene MEN1 transcribed sequence (cDNA).
XX
XX Familial multiple endocrine neoplasia type 1; PMEN1; MEN1;
KW menin; tumour suppressor gene; cancer; marker; diagnosis;
KW gene therapy; human; ss.
XX
XX Homo sapiens.
OS
XX
XX Key Location/Qualifiers
XX CDS 111..1943
XX FT /*tag= a
XX FT polyA_signal 2744..2748
XX FT /*tag= b
XX
XX WO9839439-A1.
XX
XX 11-SEP-1998.
XX
XX 04-MAR-1998; 98WO-US04258.
XX
XX 05-MAR-1997; 97US-0040269.
XX
XX (USSH ) US DEPT HEALTH & HUMAN SERVICES.
XX
XX Agarwal SK, Burns AL, Chandrasekharappa SC, Collins FS;
PI Debelenko LV, Emmert-Buck MR, Guru SC, Liotta LA;

```

PI Lubensky IA, Manickam P, Marx SJ, Spiegel AM;
 XX WPI: 1998-506360/43.
 DR P-PSDB; AAW29749.
 DR
 XX Protein and gene associated with multiple endocrine neoplasia type 1
 PT - useful in gene therapy and to diagnose sufferers of, and those
 PT susceptible to, this condition by detecting protein absence or gene
 PT mutation(s)
 XX
 XX Claim 7: Page 55-56; 75pp; English.
 PS
 XX This is the transcribed sequence (cDNA) of human MEN1, a novel
 CC tumour suppressor gene which is associated with multiple endocrine
 CC neoplasia type 1. MEN1 cDNA can be obtained e.g. from leukocyte
 CC RNA using MEN1-specific probes and PCR primers. A genomic sequence
 CC (see AY57416) is also provided. MEN1 encodes a 67.5 kDa protein
 CC (see AAW28749), termed menin, that exhibits no apparent similarities
 CC to previously known proteins. The lack of a functional menin
 CC polypeptide, either by absence of the protein, its alteration
 CC and/or associated mutations in the MEN1 gene, have been identified
 CC in individuals with familial multiple endocrine neoplasia type 1
 CC (FMEN1) and suffering from multiple endocrine neoplasia type 1.
 CC The identification of MEN1 provides a new window into the mechanism
 CC of endocrine tumorigenesis, facilitates accurate early diagnosis of
 CC MEN1 associated cancers, and provides preclinical identification of
 CC individuals with the FMEN1 syndrome, i.e. cancer-free individuals
 CC that are at high risk of acquiring MEN1 associated tumours. It
 CC also provides a basis for gene therapy. The MEN1 nucleic acids may
 CC also be used therapeutically to produce menin recombinantly and to
 CC produce transgenic animals useful in research.
 XX
 SQ Sequence 2772 BP; 577 A; 909 C; 779 G; 507 T; 0 other;

Query Match 100.0%; Score 2772; DB 19; Length 2772;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 2772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGTGTCGGAGCCGCGACCTAGAGATCCAGAGCCACAGCGCGCGCCCGCCGCC 60
 DB 1 ggttcgagc 60
 QY 61 ACTATTTCAGGCTCTCGGGGAGGGGCGCGCGCCACCGCGCGCGCATGGGGTGA 120
 DB 61 actatttcaggctctcggggagggcgccgcgcgcgcgcgcgcgcgcgcgcgcgc 120
 QY 121 AGCGCGCCAGAGACGCTGTCCGCTGCGCTCCATCGACGAGCTGCTGCGCTGTTG 180
 DB 121 agccgc 180
 QY 181 CTGCGGAGCTGGGCGGAGAGCGCGACCTGTGCTTCCTTCTGCTGGGCTTCG 240
 DB 181 ctgc 240
 QY 241 TGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTACCAAGTTCCCGAGCTCACTTC 300
 DB 241 tggagcattttctggctgtcaacgcgctcatccctacaaagttcccgagctcacttc 300
 QY 301 AGCCGAGCCCGCCGCGCGCTGGCGGCTCCTACTTTCCTGCGCGCACCTGT 360
 DB 301 agccgagcccgccgcgcgctggcggcctcacttctcccgctggcgacgtgt 360
 QY 361 CTATCATCTGGCGGCTCTATGCGCGCTTACCGCGCAGATCCGAGCGCGCTCGACCTGT 420
 DB 361 ctatcatctggcggcctctatgcgcgcttaccgcgcgcgcgcgcgcgcgcgcgcgc 420
 QY 421 CCCTCTATCTCGAGAGGGGGTGTCTCCACCGCTGAGCTGGTGAAGAAGTCTCCGATG 480
 DB 421 ccctctatctcgagaagggtgtctccagcgtgagctggtgaagaaggtctccgatg 480
 QY 481 TCATATGAACAGCCTCAGCGCTCTACTTCAAGGATCGGGCCACATCCAGTCCCTCT 540
 DB 481 tcatatggaacagcctcagcgctctacttcaaggatcgggccacatccagttccctct 540

DB 481 tcatatggaacagcctcagcgcgtctctacttcaaggatcgggccacatccagtcctct 540
 QY 541 TCAGCTTCATCAGAGCACCACAAATTGGACAGCTCCGGTGTGGCTTTGCTGTGTGGG 600
 DB 541 tcagcttcatacagagcaccacaaattggacagctccgggtgtggccttctgctggttgg 600
 QY 601 CTGCGCAGGCGCTGGGTCTCCGGGATGTCCACCTCGCCCTGTGTGAGGATCATGCTTGG 660
 DB 601 ctgcgagcgccctgggtctccgggatgtccacctgcacctgtctgaggatcatgacctgg 660
 QY 661 TAGTGTTTGGGCCCCAATGGGGAGCAGACAGCTGAGGTTCACCTGCGCAGCAAGGCAAGC 720
 DB 661 tagtgtttggcccacaaatggggagcagacagctgaggtcacctggcagcggaaggcaag 720
 QY 721 AGACCGCAGCGGCGCAGACAGTCAATGCCGCTGTGGCTGAGCGGAGTGTGCTGTACCTGA 780
 DB 721 agaacgcagggcgccagacagtcacgctgggtgtgctgagcagagctgctgtacctga 780
 QY 781 AAGGATCATACATGCGCTGTGACCGCAAGATGAGGTGGCGTTTCATGGTGTGTGCCATCA 840
 DB 781 aaggatcatacatgctgtgacccgaagatggaggtggcgcttcatgtgtgtgcatca 840
 QY 841 ACCCTTCATTTGACCTGCACACGACTCGCTGGAGCTTCTGCAGCTGCACGAGAAGCTGC 900
 DB 841 acccttcatactgacctgcacacgcagctcgctggagcttctgcaagctgcaagaagctgc 900
 QY 901 TCTGGCTGCTCTATGACCTGGGACATCTGGAAGGTACCCCATGGCTTATAGGAAACCTGG 960
 DB 901 tctggctgctctatgacctggagctggacatctggaaggtaacctatgaggaacctgg 960
 QY 961 CAGATCTAGAGGAGCTGGAGCCACCCCTGGCGGCGCAGACCCACTCACCTCTACCAACA 1020
 DB 961 cagatctagaggagctggagccacacctggcgggccagacccactcacctctaccaca 1020
 QY 1021 AGGCATTTGCTCAGCAAGACCTACTATCGGATGAACACATCTACCCCTTACATGTACC 1080
 DB 1021 aggcattgcttcagccaagacctactatctcggatgaacacatctacctctacatgtacc 1080
 QY 1081 TGCGTGSCTACCACTGTGCAACCGCAATGTGCGGGAAGCCCTGCAGGCTTGGCGGACA 1140
 DB 1081 tggctgctaccactgtcgcaacgcgaatgtgcgggaagccctgcaggcctggcggaaca 1140
 QY 1141 CGGCCACTGTCTCCAGGACTACAACCTACTTCCCGGGAAGAGGAGAGATCTACAGGAGT 1200
 DB 1141 cgccactgtctcagagactacaactactgcggggaagcagagagatctacaaggagt 1200
 QY 1201 TCTTTGAAGTAGCAATGATGTATCCCAACCTGCTGAAGGAGGCGACGCTTGTCTGG 1260
 DB 1201 tctttgaagtagccaatgatagtcatccccaacctgtcgaaaggagcagcagcttgcctgg 1260
 QY 1261 AGCGGCGGAGGAGCGCGCGGAGCAAGCCAGCGCACCCAGAGCCAAGTTTCGCGCC 1320
 DB 1261 agcgcgcgagagcgcgcggggagcaagcagggcaacccagagccaagggttccgccc 1320
 QY 1321 TCAGGACCTGTAGTGTCTGCCACCTGTGCGATTCAGGAGGATCTGCAAAATGGG 1380
 DB 1321 tcaggacacctgagtgttcgccacctgtgcgattctacgcggtcatctgcgaatggg 1380
 QY 1381 AGGAGGCGAGTCCACGCTGTGTCACGCTGGGCTGGGCGCACCTTTCTGTGAGTCCC 1440
 DB 1381 agagggcagtcaccaacctgtgctgcagctggcgaccttcttctgtgcagctccc 1440
 QY 1441 TAGGCCCTTTTGGAGGACAGTGTGGCGCAGAGGTGCGCATAGTAGCCGAGAGCCGAGG 1500
 DB 1441 taggcccttttggagacaggtgcgcgagaggtgcgcatagtgagccgagagccgaggg 1500
 QY 1501 CGGCGGAGCGGAGGACCGTGGGGCGAGGAAGCCCGGGAGGCGCGCGGGGGGGCCAC 1560
 DB 1501 cggcgagggcgagagcggtggggcgaggaagcccggaagcggcgcgcgggggccac 1560
 QY 1561 GCGGGGAGTCCAAGCCAGAGAGCGCCCGCGCCCAAGAGCCAGCACTGGACAAAGGCC 1620
 DB 1561 ggcggggagtcgaagccag 1620

```
QY 1621 TGGGACCCGCGCAGGGTGCAGTGTTCAGGACCCCGGAGGCTCTCTGGACTGTGCGTGG 1680
|||||
Db 1621 tgggcaacgcgccaggggtgcagtgtcaggaccgcccggaagcctcctggagactgtcgctg 1680
QY 1681 GCACAGCCGCGAGGCGCTGAAGTGTGCAGACACGGCTCAGGTGCAGACCCGCGAGCATCAC 1740
|||||
Db 1681 gcacagcccgagggcctgaagtggcagcagcgctcaggtgcagcaccgcgagcatcac 1740
QY 1741 CACCGCCGAGGGTCCAGTGTCTCACTTCCAGAGTCAGAGATGAAGGCATGAAGGAGC 1800
|||||
Db 1741 caccgcccggaggggtccagtgctcaacttccagagtgagaagatgaaggagc 1800
QY 1801 TGCTGTGGCCACCAAGATCAACTCGAGCGCATCAAGCTGCAACTCAGCGCACAGTCCG 1860
|||||
Db 1801 tgctgtggcccaacgaatcaactcagcgccatcaagctcgaactcagcgccagctcgc 1860
QY 1861 AAGTCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTGTACTACACTCTGTCTTCTCTCA 1920
|||||
Db 1861 aagtgcatgaagaagcagaagtgccaccccttagtgactacactctgtctttcctca 1920
QY 1921 ACGCGCAGCGCAAGGCTCTCAACTACTGTGGGACTTCGGACCGCTGTGGGGACCCAGG 1980
|||||
Db 1921 agcgcagcgcaaaagcctctgaactctggaacttgggagcttggtgggacccagg 1980
QY 1981 CTCGCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGC 2040
|||||
Db 1981 ctccgcttagtccccaactcagcccatgtctgccccagcccaaaaggggacagcg 2040
QY 2041 CTCACCTCTACCCCAACCCCTAGTTCGCGGTCCCGAGTACAGTCTGTATCAACCCACGA 2100
|||||
Db 2041 ctcaactctaccacaacccctaggttcccggtcccgagtcacagtcgtatcaacaaccaa 2100
QY 2101 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCGAGTCTGTTAGAAATATAGGTCTCTCTCC 2160
|||||
Db 2101 ttcttcagctcagaaacccagggctctgccccagctcgtagaataggtctctctcc 2160
QY 2161 CAGAATCCAGCGCGCCCAATGTAACCTCACGCTGGTCTTAATACCAAGTCTTTAAAGG 2220
|||||
Db 2161 cagaatccagcgcccaatgtaacctcaactcagctggctccttaattaccagttttaagg 2220
QY 2221 CCGAGCCCTAGAACCCAGCTCTCTCTCGGAACCGCTCAGTACAGCCAGCAACAGT 2280
|||||
Db 2221 ccagccccctagaaaccccaagctctctctcggaacccgctcactcagagccagcaaacg 2280
QY 2281 TACTCAGGCTCTCTCCAGCTGTAGGAGCTCAGGTTTCACCCCTTAACCAAGGAGGAC 2340
|||||
Db 2281 tactcagggctctctccagcttctgaggagctgaggtttcacccctlaacccaaggagcac 2340
QY 2341 AGGTCCCACTCCAGCGCGGAGGCTTAGGACCACTAGCCCTTAGGAGTATATTTCCGC 2400
|||||
Db 2341 aggtcccaactccagccgggagcctaggaccactcagccctaggagatatattccgc 2400
QY 2401 ACTTCAGATTCCTATCTTGGCAATCCAGTCCCTGCCCCCAATAATCTCAGTCTCTGC 2460
|||||
Db 2401 acttcagaattccatctcttgcgaatccaaagctcctgcccccaataaacttcagtcctgc 2460
QY 2461 TTCCAGAAATTTGGAATCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2520
|||||
Db 2461 ttccagaatttggaaatctctctctctctctctctctctctctctctctctctctctctct 2520
QY 2521 TCCGCCCCCAGGCTATGAGCATCTGAGCCCGCCCTCTTCTCTGAGCAACTGGCCCCCGG 2580
|||||
Db 2521 tcgccccccagcctatgagcatcctgagcccgccctctctctctgagcaaacctggccccg 2580
QY 2581 ATCAGAGGAGGACTCTCTTCGACCTCTGGAACTCCGAGTCCCGAGGTCGAGCCCATCTCG 2640
|||||
Db 2581 atcagagggagcctctctctccgacctctggaaacctcccgaggtccagcccatctcg 2640
QY 2641 GAGCATCCCGAGGAAATCTGAGGGGCTTAGGAGTGGGTGACAGAGCCCTGATCTCTT 2700
|||||
Db 2641 gagcatcccgaggaaatctgagagggttaggagtggtgacaagagccctgactctt 2700
```

```
QY 2701 CCTGTTTGTACATGATTTATTTTTCAGTCCAGAAAAGATCAATFACATTTGTTAAAA 2760
|||||
Db 2701 cctgtttgtacatgatttttttttccagttccaaagaagatgaatacatatttgttaaaa 2760
QY 2761 AAAAAAATAA 2772
|||||
Db 2761 aaaaaaataa 2772

RESULT 2
AAV80659
ID AAV80659 standard; cDNA; 2196 BP.
XX
AC AAV80659;
XX
DT 15-MAR-1999 (first entry)
XX
DE Human multiple endocrine receptor neoplasia type 1 protein cDNA.
XX
KW Human; multiple endocrine neoplasia type 1 protein; MEND-1;
KW tumorigenesis; cancer; multiple endocrine neoplasia; gene therapy;
KW detection; diagnosis; drug screening; ss.
XX
OS Homo sapiens.
XX
FH Key
CDS
FT
FT /tag= a
FT /product= "multiple endocrine neoplasia type 1 protein"
FT /transl_except= (pos:466..468,aa:Xaa)
FT /transl_except= (pos:538..540,aa:Xaa)
FT /transl_except= (pos:877..879,aa:Xaa)
FT /transl_except= (pos:991..993,aa:Xaa)
FT /transl_except= (pos:1306..1308,aa:Xaa)
FT /transl_except= (pos:1318..1320,aa:Xaa)
FT /transl_except= (pos:1327..1329,aa:Xaa)
FT /transl_except= (pos:1339..1341,aa:Xaa)
FT /transl_except= (pos:1360..1362,aa:Xaa)
FT /transl_except= (pos:1384..1386,aa:Xaa)
FT /note= "Xaa is unspecified"
XX
PN W09854324-A1.
XX
PD 03-DEC-1998.
XX
PF 29-MAY-1998; 98WO-US10957.
XX
PR 29-MAY-1997; 97US-0865337.
XX
PA (INCY-) INCYTE PHARM INC.
XX
PI Au-Young J, Covitz PA, Murry LE, Tang YT;
XX
DR WPI; 1999-059839/05.
DR P-PSDB; AAW86348.
XX
PT New isolated multiple endocrine neoplasia type 1 protein - used to
PT develop products for the diagnosis, treatment and prevention of
PT cancers and multiple endocrine neoplasia.
XX
PS Claim 5; Fig 1; 67pp; English.
XX
CC The present sequence encodes human multiple endocrine neoplasia type 1
CC protein (MEND-1). MEND-1 plays a role in multiple endocrine neoplasia
CC when one or both normal genetic copies of MEND-1 are mutated and no
CC longer able to suppress tumorigenesis. MEND-1 can be used for treating
CC or preventing cancers and multiple endocrine neoplasia. MEND-1
CC polynucleotides can also be used for gene therapy; products from the
CC present invention can also be used for detection, diagnosis and drug
CC screening.
XX
SQ Sequence 2196 BP; 445 A; 712 C; 623 G; 385 T; 31 other;
```

[illegible]

Qy	1108	ATGTGCGGGGAAGCCCTGCAGGCGCTGGG----	CGGACACAGGCCACTCTTCATCCAGGACATACA	1164
Db	980	: : :~	: : :~	1039
		atgtcgagggaascccttgcagcggcccttgcagagacgagcaacgacctcttcaccagactaca		
Qy	1165	ACTACTGCGGGGAAGACGAGGAGATCTCAAGGAGTCTTTTGAACTACGCCAATGATGCA	1224	
Db	1040	actacttcggygaagacgagagatactaagaagttctttgaagttagccaattgatgtca	1099	
Qy	1225	TCCCCAACCTTGCTGAAGGAGCAGCACGCTTGCGAGGCGGCGGAGGAGCGCGCGGGGG	1284	
Db	1100	tccccaaaccttgctgaaggaaggaagcagcagcttcttggaagggcgagagagcgagcgagggg	1159	
Qy	1285	AGCAAAGCCAGGCGACCACAGAGCCAAGTTCCGCCCTCCAGGACCCTGAGTGCTTCGCCC	1344	
Db	1160	agcaaagccagggcaccccagagccaaggttccgcctccaggaccctgagtgcttcgcc	1219	
Qy	1345	ACCTGTGCGCATTTCTAGGACGGCATCTGCAATGGGAGGAGGGCAGTCCCAGCGCTGTGC	1404	
Db	1220	accttgtgcattctctacgacggcatctgcaaatggagagggcgagtcgcccgctgtgc	1279	
Qy	1405	TGCAGTGGGCTGGGCGACCTTTCTTGTCAGTGCCTTAGCGCGCTTTTGAGGACACAGGTGC	1464	
Db	1280	tgcactggyggtctgggacaccttctctgncagtcctctangccgttttganggaaggtgc	1339	
Qy	1465	GGCAGAAGTGTGCGCATAGTGAGCGGAGAGCGGCGCGGCGGCGGAGCGGAGGCGGTGGG	1524	
Db	1340	ngcagaaggttgcgcatagtgancagagatlgcgcgaggtctgc-----	1380	
Qy	1525	GCGAGGAAGCCCGGAGAGCGCGCGCGGCGGCGGCGGAGGTGCANATGCAGAGGAGC	1584	
Db	1381	-----gagncacgaagccagaggagc	1399	
Qy	1585	CCCGCGCGCCAAAGACCGAGCATTGGCAAGGGCTTGGGCACCGCGCCAGGCTGCAGTGT	1644	
Db	1400	ccccgcgcgccaagagccagcaacttggaagaaggccttggccacgcggccagggcgagtg	1459	
Qy	1645	CAGGACCGCCCGGAAGCCCTCTTGGGACATGTGCGTGGACACAGCCGCGAGCGCCCTGAAGGTG	1704	
Db	1460	caggaaaccccccggaagcctctgggaacttgctgtgcgtggcaagccgagccctgaaggtg	1519	
Qy	1705	GCAGCAGCGGTCTAGGTGCCAGCACCCCGCAGCATATCACACC CGCGGAGGGTCCAGTGTCTN	1764	
Db	1520	gcagcagcgtctcaggttgccagcaaccgcgcagcatcaccacgcgcgaggggtccagtgctca	1579	
Qy	1765	CTTTCACAGCTGACAGATTAAGGGCATCAAGGAGCTGCTGTGTGCCACACCAAGATCAACT	1824	
Db	1580	ctttccagagtgaagagatgaagggcatgaaaggagctgctgtgtgcccaccaaga tcaact	1639	
Qy	1825	CGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCAAGTGCAGATGAAGAGCAGAAG	1884	
Db	1640	cgagcgcca tcaagctlgcaactcacgcgcacagctgcgaagtgcagatgaagaagcagaag	1699	
Qy	1885	TGTCACCCCTAGTGACTACACTCTCTTTTCTCAAGCGGCGAGCGCAAGAGCCTCTGAA	1944	
Db	1700	tgtcacacccctagtgtactaacctctcttcttccaaagcgcgaggaagagcctctgtga	1759	
Qy	1945	CTACTGGGGGACTTCGAGCCGCTTGTGGGAGCCAGGCTCCG--CCTTACTGTCGCCCAACTCT	2003	
Db	1760	ctactggyggacttcgagccgcttgtygggaaccaggtctcgcccttagtcccccaactct	1819	
Qy	2004	GAGCCCATGTCTTGCCCGCCAGCCCAAGGGGACAGGCGCTCACCTCTTACCCAAACCTTAGG	2063	
Db	1820	gagccaatgttctgcccccaagccccaaagggagcagggcctcaactctacccaaaaccttag	1879	
Qy	2064	TTCCCGGTCCCGAGTACAGTCTGTATCAACGCCACGATTTTCTCGAGCTCAGAACCCAGG	2123	
Db	1880	ttcccggtccccgagtacagtctgtatcctgaaccccaagattttctecagctcagaaacccag	1939	
Qy	2124	GCTGTGCCCCAGTCTAGATAATAGTCTCTTCCTCCAGATCCACCGCGCCCAATCGA	2183	
Db	1940	gctctgccccagtcgttagaataataggtctcttctccccagatatccagcccgagmcatgga	1999	
Qy	2184	AACCTCAGCTGGGTCCCTAATTACCACTCTTTAAAGCGCCAGCGCCCTAGAAACCCCAAGCT	2243	

FT		/number= 7	
FT	exon	6623..6758	
FT		/*tag= p	
FT		/number= 8	
FT	intron	6759..7195	
FT		/*tag= q	
FT		/number= 8	
FT	exon	7196..7360	
FT		/*tag= r	
FT		/number= 9	
FT	intron	7361..7577	
FT		/*tag= s	
FT		/number= 9	
FT	exon	7578..9180	
FT		/*tag= t	
FT		/number= 10	
FT	TATA_signal	8861..8865	
FT		/*tag= e	
XX			
PN	WO9839439-A1.		
XX			
XX			
PD	11-SEP-1998.		
XX			
Pf	04-MAR-1998;	98WO-US04258.	
XX			
PR	05-MAR-1997;	97US-0040269.	
XX			
PA	(USSH) US DEPT HEALTH & HUMAN SERVICES.		
PI	Agarwal SK, Burns AL, Chandrasekharappa SC, Collins FS;		
PI	Debelenko LV, Emmert-Buck MR, Guru SC, Liotta LA;		
PI	Lubensky IA, Manickam P, Marx SJ, Spiegel AM;		
DR			
XX	WPI: 1998-506360/43.		
DR	P-PSDB; AAW29749.		
XX			
PT	Protein and gene associated with multiple endocrine neoplasia type 1		
PT	- useful in gene therapy and to diagnose sufferers of, and those		
PT	susceptible to, this condition by detecting protein absence or gene		
PT	mutation(s)		
XX			
PS	Claim 32; Page 57-62; 75pp; English.		
XX			
CC	This is the nucleotide sequence of human MEN1, a novel tumour		
CC	suppressor gene which is associated with multiple endocrine		
CC	neoplasia type 1. MEN1 genomic DNA can be obtained from genomic		
CC	libraries using MEN1-specific probes and PCR primers. A cDNA clone		
CC	(see AAU57415) is also provided. MEN1 encodes a 67.5 kDa protein		
CC	(see AAU28749), termed menin, that exhibits no apparent similarities		
CC	to previously known proteins. The lack of a functional menin		
CC	and/or associated mutations in the MEN1 gene, its alteration		
CC	in individuals with familial multiple endocrine neoplasia type 1		
CC	(FEMEN1) and suffering from multiple endocrine neoplasia type 1.		
CC	The identification of MEN1 provides a new window into the mechanism		
CC	of endocrine tumorigenesis, facilitates accurate early diagnosis of		
CC	MEN1 associated cancers, and provides preclinical identification of		
CC	individuals with the FEMEN1 syndrome, i.e. cancer-free individuals		
CC	that are at high risk of acquiring MEN1 associated tumours. It		
CC	also provides a basis for gene therapy. The MEN1 nucleic acids may		
CC	also be used therapeutically to produce menin recombinantly and to		
CC	produce transgenic animals useful in research.		
XX			
SQ	Sequence 9180 BP; 1936 A; 2604 C; 2612 G; 2027 T; 1 other;		

Query Match	47.3%;	Score 1312.2;	DB 19;	Length 9180;
Best Local Similarity	96.5%;	Pred. No. 3.2e-283;		
Matches 1341;	Conservative 0;	Mismatches 48;	Indels 0;	Gaps

Qy	1382	GGAGGCGAGTCCCACGCCTGTGTGTCACGTGGGTGGGCCACCCTTTCTTGTCAGTCCCT	144
Db	7499	gqatggcagacgaaggcttcgaagtccagaccactagccgcaccttgctccacctt	755

XX	AAV80662;
AC	
XX	
DT	15-MAR-1999 (first entry)
XX	
DE	Human multiple endocrine receptor neoplasia type 1 clone 1313372.
XX	
KW	Human; multiple endocrine neoplasia type 1 protein; MEND-1;
KW	tumorigenesis; cancer; multiple endocrine neoplasia; gene therapy;
KW	detection; diagnosis; drug screening; ss.
OS	Homo sapiens.
XX	
PN	WO9854324-A1.
XX	
PD	03-DEC-1998.
XX	
Pf	29-MAY-1998; 98WO-US10957.
PR	29-MAY-1997; 97US-0865337.
XX	
PA	(INCY-) INCYTE PHARM INC.
PI	Au-Young J, Covitz PA, Murry LE, Tang YT;
XX	WPI; 1999-059839/05.
DR	
XX	
PT	New isolated multiple endocrine neoplasia type 1 protein - used to
PT	develop products for the diagnosis, treatment and prevention of
PT	cancers and multiple endocrine neoplasia.
XX	
PS	Disclosure; Page 49; 67pp; English.
XX	
CC	The present sequence represents Incyte clone 1313372 which is used to
CC	make up the consensus human multiple endocrine neoplasia type 1 protein
CC	(MEND-1) (see AAV80659). MEND-1 plays a role in multiple endocrine
CC	neoplasia when one or both normal genetic copies of MEND-1 are mutated
CC	and no longer able to suppress tumourigenesis. MEND-1 can be used for
CC	treating or preventing cancers and multiple endocrine neoplasia. MEND-
CC	polynucleotides can also be used for gene therapy. Products from the
CC	present invention can also be used for detection, diagnosis and drug
CC	screening.
XX	
SQ	Sequence 812 BP; 188 A; 270 C; 209 G; 131 T; 14 other;
Query Match	26.9%; Score 746.4; DB 20; Length 812;
Best Local Similarity	97.2%; Pred. No. 2.5e-157;
Matches 772; Conservative	4; Mismatches 16; Indels 2; Gaps
QY	1570 CCAAGCGCAGAGGACGCCCGCCGCCCAAGAAGCGACTGGACAAGGGCTGTGGCACCG 1629
Db	
Db	1 ccaagccagaggagccccgcgccaaagaagcagactggacaaaggccctgggaccag 60
QY	1630 GCCAGGGTGCAGTGTCAGGACCCCCCGGAAGCCCTCTCGGACTGTCTCGCTGCCACAGCCC 1689
Db	61 gccaggcgcatgtcaggaaccccccgaaaacctcttgggaactgtcgatggacagccc 120
QY	1690 GAGGCCCTGAAGTGGCAGCAGCGGTCTAAGTGCCAGCACACC CGCAGCATACACCGCGGG 1749
Db	121 gaggccctgaagtggcagcacggttcaggtgccagccccgcgcagcatcacccagccg 180
QY	1750 AGGGTCCAGTGCTCATTTCAGAGTGAGAAATGAAGGGCATGAAGCAGCTGTCTGTGGTGG 1809
Db	181 agggctcagtgctcactttccagagtgagaagatgaaggcatgaaggagctgctgtg 240
QY	1810 CCACCAAGATCAACTCGAGCGGCATCAAGCTCCAACCTCAGGCACAGTCGCAAGTGCAGA 1869
Db	241 ccaccaagataccaactcgagcgccatacaagtgcactcacgagcagtcgcaagtcgaga 300
QY	1870 TGAAGAGCAGAGAAGTGTCACCCCTAGTAGTACTACACTCTGTCTTTTCTCAAGCGGCAGC 1929
Db	301 tgaagaagcagagaagtgtccacccctagtagctacactctatcttctctcaagcgcag 360

PD	01-JUN-1995.	
XX		
XX	11-NOV-1994;	94WO-JP01916.
XX		
XX	12-NOV-1993;	93JP-0355504.
XX		
XX	(MATS/) MATSUBARA K.	
PA	(OKUB/) OKUBO K.	
PA		
XX		
PI	Matsubara K, Okubo K;	
XX		
DR	WPI; 1995-206931/27.	
XX		
PT	Identifying gene signatures in 3'-directed human cDNA library - e.g.	
PT	for diagnosis of abnormal cell function, by preparing cDNA that	
PT	reflects relative abundance of corresp. mRNA in specific human	
PT	tissues	
XX		
XX	Claim 1; Page 1227; 2245pp; Japanese.	
XX		
CC	A single-stranded DNA (or its complementary strand or the corresp.	
CC	double-stranded DNA) which comprises one of the 7837 "GS" sequences	
CC	given in AAT19001-T26837 and which is able to hybridise to part of	
CC	human genomic DNA, cDNA or mRNA is claimed. The GS (Gene Signature)	
CC	sequences were obtained from 3'-directed cDNA libraries prepared	
CC	from various human tissues; synthesis of cDNA was initiated from the	
CC	3'-end of mRNA by using poly(T) as the sole primer. Since the 3'-	
CC	untranslated sequence is unique to a particular mRNA species, almost	
CC	all the 3'-oriented cDNAs hybridise with specific mRNAs. Each library	
CC	is constructed so as to reflect accurately the relative abundance of	
CC	different mRNAs in the particular tissue from which it was derived.	
CC	The appearance frequency of a given GS in a cDNA library can be	
CC	determined (esp. using primers and probes derived from the GS	
CC	sequences) as a means of diagnosing abnormal cell function or for	
CC	recognising different cell types.	
XX		
SQ	Sequence 67 BP; 20 A; 9 C; 9 G; 28 T; 1 other;	
	Query Match	2.4%; Score 66; DB 16; Length 67;
	Best Local Similarity	98.5%; Pred. No. 6,7e-06;
	Matches 66; Conservative	0; Mismatches 1; Indels 0; Gaps 0;
QY	2693 GATCTCTTCCTGTTTCFACATAGATTATTTTTCAGTCCACGAAAGATGATACATT	2752
Db		
	1 gatctcttcctgtttcacatagattatttttcagttcccaagaagatgacacatt	60
QY	2753 TGGTTAAA	2759
Db		
	61 tgtnaaa	67
RESULT	9	
AAZ87297		
ID	AAZ87297 standard; DNA; 13842 BP.	
XX		
AC	AAZ87297;	
XX		
DT	05-JUN-2000 (first entry)	
XX		
DE	S. venezuelae macrolide biosynthetic gene pikAI, SEQ ID NO:30.	
XX		
KW	Desosamine biosynthesis; macrolide; polyketide; methymycin; pikromycin;	
KW	neomethymycin; narbomycin; polyhydroxyalkanoate monomer synthase;	
KW	biopolymer; antibiotic; chemotherapeutic; immunosuppressant; asthma,	
KW	chronic obstructive pulmonary disease; respiratory inflammation;	
KW	hypercholesterolaemia; crop protection agent; ds.	
XX		
OS	Streptomyces venezuelae ATCC15439.	
XX		
Key	Location/Qualifiers	
CDS	1..13842	
FT	/*tag=	
PT		

monomers. The compounds produced by the recombinant host cells are useful as biopolymers, e.g., in packaging or biomedical applications, to engineer PBA monomer syntheses or to prepare biologically active agents, such as chemotherapeutics, immunosuppressants, agents to treat asthma, chronic obstructive pulmonary disease as well as other diseases involving respiratory inflammation, cholesterol-lowering agents or macrolide-based antibiotics which are active against a variety of organisms, e.g., bacteria, including multi-drug resistant pneumococci and other respiratory pathogens, as well as viral parasitic pathogens, or as crop protection agents (e.g., fungicides or insecticides) via expression of polyketides in plants. The present sequence represents the macrolide biosynthetic gene cluster (pk) from *Streptomyces venezuelae* ATCC 15439.

[illegible]

RESULT	12
AAA75633	
ID	AAA75633 standard; DNA; 38506 BP.
XX	
XX	AAA75633;
XX	
DT	22-JAN-2001 (first entry)

DE	Nucleotide sequence of the insert DNA in cosmid pKOS023-27.
DX	
KW	Narbonolide synthase; polyketide synthase gene; narbonolide polyketide;
KX	antibiotic; C12-hydroxylase; pick; desosamine biosynthesis;
KW	desosaminyl transferase enzyme; ketolide; beta-glucosidase enzyme;
KW	picromycin biosynthesis; ss.
XX	
OS	Streptomyces venezuelae.
XX	
US	US6117659-A.
PN	
XX	
PD	12-SEP-2000.
XX	
PF	27-MAY-1999; 99US-0320878.
XX	
PR	28-MAY-1998; 98US-0087080.
PR	22-SEP-1998; 98US-0100880.
PR	08-FEB-1999; 99US-0119139.
PR	20-MAY-1999; 99US-0134990.
PR	30-APR-1997; 97US-0846247.
PR	06-MAY-1998; 98US-0073538.
PR	28-AUG-1998; 98US-0141908.
XX	
XX	(KOSA-) KOSAN BIOSCIENCES INC.
XX	
PI	Ashley G, Betlach MC, Betlach M, Tang L, Mcdaniel R;
XX	
WI	PI; 2000-610844/58.
XX	
PT	New recombinant pick hydroxylase gene of Streptomyces venezuelae useful
PT	for converting ketolides to antibiotics and as antibiotics and
PT	intermediates in the synthesis of compounds with pharmaceutical value
PT	-
XX	
PS	Disclosure: Columns 15-32; 117pp; English.
XX	
CC	The present sequence is used to produce the recombinant DNA compounds
CC	of the invention. The specification describes a recombinant DNA compound
CC	expressing recombinant polyketide synthase genes in host cells for the
CC	production of narbonolide, narbonolide derivatives and polyketides that
CC	are useful as antibiotics and as intermediates in the synthesis of
CC	compounds with pharmaceutical value. The DNA compounds may also encode
CC	a C12-hydroxylase (pick), desosamine biosynthesis and desosaminyl
CC	transferase enzymes (useful for conversion of ketolides to antibiotics)
CC	and the beta-glucosidase enzyme (involved in picromycin biosynthesis).
CC	These compounds are also useful for increasing the antibiotic activity
CC	of a compound relative to the unhydroxylated compound. The recombinant
CC	host cells are useful as genetic systems that allow rapid engineering
CC	of the narbonolide polyketide synthase. These would be valuable for
CC	creating novel ketolide analogs for pharmaceutical applications.
XX	
SQ	Sequence 38506 BP; 4914 A; 15118 C; 13444 G; 5030 T; 0 other;

	Query Match	1.9%	Score 52.6;	DB 21;	Length 38506;
	Best Local Similarity	45.3%;	Pred. No. 0.052;		
Matches	274;	Conservative	0;	Mismatches 324;	Indels 7; Gaps 2;
Qy	52	CGGGCCGCCACTATTTCCAGGCTCTCGGGGCAGGGCGCGCGCC-CACCGCCCGCGCCGCC	110		
Db	9676	cggccgcccgcactgtctcggtcaactcgttcggagagatcgccgcgcgcacgttcgcgcgtt	9735		
Qy	111	ATGGGGGTGAAGGCCGCCCAAGAAGACGCTGTTCOCGTGGTCCATCGACAGCTGGTG	170		
Db	9736	gtgtttcgtctgcgccgacgcgcgcctgtgctgcgcgcgcgcgtcatgcaggag	9795		
Qy	171	CGCCTGTTTCTGCGGAGCTGGGGCGAGAGAGCGGACCTGGTGCTCTCTTCTTGTGTG	230		
Db	9796	ctgcgccgcggttgagcgagatctgcgcgtccaggccgcgaggacagatccgcgtctg	9855		
Qy	231	CTGGGCTTCGTGGAGCATTTTTCTGGCTGTCAACCGCGTATCCCTACCACGTTCCCGAG	290		
Db	9856	ctgagacgaagaacgaatacacggagacgtctgagatctgcgcacctcaaacgccccaq	9915		

XX	key	Location/Qualifiers
PH	CDS	20.13912
FT		/*tag= a
FT	CDS	14056.14136
FT		/*tag= b
FT	CDS	14148.15827
FT		/*tag= c
XX		
PN	WO9722711-Al.	
XX		
PD	26-JUN-1997.	
XX		
XX	18-DEC-1996;	96WO-US20119.
PF		
XX		
XX	19-DEC-1995;	95US-0008847.
PR		
XX	(MINU) UNIV MINNESOTA.	
PA		
XX	Sherman DH, Williams MD, Xue Y;	
PI		
XX	WPI; 1997-341701/31.	
XX	P-PSDB; AAW19629-30 AND AAW00918.	
DR		
DR		

Query Match		1.9%	Score 52.2;	DB 18;	Length 15872;
Best Local Similarity		44.9%;	Pred. No. 0.048;		
Matches 196;	Conservative	0;	Mismatches 243;	Indels	0; Gaps
0;	Gaps				
Qy	44	CAGCGGCGGGCCGCCACATATTCCAGGCTCTGCGGGGCAGGGCGCGGCCACACGCC	103		
Db	4729				
		cggcgaaaccgcccagggccgcacgaaggccttcgtttcacgccggaaggagcgacgg	4788		
Qy	104	CGCGCGCATGGGGTGAAGGCGGGCCAGAACACGCTGTTCGCGTGCCTGCATCGACGA	163		
Db	4789				
		cgtcgccatggcgaggaactcgccgcgcacacctgttgcgccgcgcctcgacac	4848		
Qy	164	CGTGCGGCGCTGTTGCTGCGAGCTGGCGCGAGAGAGCGGACCTGTGTGCTCCTTTC	223		
Db	4849				
		cgtgtacggccctcgacctcaectcgacggcgctgcgggagatcgctgcgcgcgg	4908		
Qy	224	CTTGCTGTGGCTTCGGTAGCATTTTCTGGCTGTCAACCGCGTCATCCCTACCAAGT	283		
Db	4909				
		ggaggagtgcacctaaccgctaacaccaaccgcgcctctcttgcccttcgaggtggcgct	4968		
Qy	284	TCCGAGCTACCTTCCAAGCCAGCCCGCCCGCACCGCGCTGSGCGCTCACCTACTT	343		
Db	4969				
		gttcgcctctcgaacaacacgacctgctcccgacctgctcaaccggcaactcogtcgg	5028		
Qy	344	TCCGCTGGCGAAGCTGTATATCATGCGCGCGCCTCTATGCCGCTTCACGCCGAGATCGG	403		
Db	5029				
		cgatatcgccgcgcgcactcgccgggtgctctctccctcgacgacgcgcactcgt	5088		

[illegible]

Search completed: October 14, 2001, 04:05:21
Job time: 25441 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 13, 2001, 19:53:04 ; Search time 15359.2 seconds
(without alignments)
2791.583 Million cell updates/sec

Title: US-09-380-337-1
Perfect score: 2772
Sequence: 1 GGTCGCGAGCGCGACC.....TGTAAAAA..... 2772

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 773387458 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htg_hum1:*
23: em_htg_hum2:*
24: em_htg_hum3:*
25: em_htg_hum4:*
26: em_htg_hum5:*
27: em_htg_hum6:*
28: em_htg_hum7:*
29: em_htg_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vl:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vl1:*
59: gb_vl2:*
60: gb_vl3:*
61: gb_vl4:*
62: gb_vl5:*
63: gb_vl6:*
64: gb_vl7:*
65: gb_vl8:*
66: gb_vl9:*
67: gb_vl10:*
68: gb_vl11:*
69: gb_vl12:*
70: gb_vl13:*
71: gb_vl14:*
72: gb_vl15:*
73: gb_vl16:*
74: gb_vl17:*
75: gb_vl18:*
76: gb_vl19:*
77: gb_vl20:*
78: gb_vl21:*
79: gb_vl22:*
80: gb_vl23:*
81: gb_vl24:*
82: gb_vl25:*
83: gb_vl26:*
84: gb_vl27:*
85: gb_vl28:*
86: gb_vl29:*
87: gb_vl30:*
88: gb_vl31:*
89: gb_vl32:*
90: gb_vl33:*
91: gb_vl34:*
92: gb_vl35:*
93: gb_vl36:*
94: gb_vl37:*
95: gb_vl38:*
96: gb_vl39:*
97: gb_vl40:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2772	100.0	2772	97 HSU93236	U93236 Human menin
2	2406	86.8	2633	91 BC002544	BC002544 Homo sapi
3	1865	2	67.3	9 AR081879	AR081879 Sequence
4	1781	64.2	1837	97 HSURE	Y12338 H.sapiens m
5	1648	6	59.5	3067 94 AB023400	AB023400 Rattus no
6	1622	58.5	1691	91 BC002664	BC002664 Homo sapi
7	1616	4	58.3	2984 94 AF109389	AF109389 Mus muscu
8	1616	58.3	2623	94 AF016398	AF016398 Mus muscu

9	1608.4	58.0	2947	94	AF130368	Mus muscu	
10	1585	57.2	2577	94	AF072755	Mus muscu	
11	1522	54.9	2261	94	AF130369	Rattus no	
12	1520.6	54.9	1864	94	AB023401	Mus muscu	
13	1518.2	54.8	1906	94	AF130370	Rattus no	
14	1312.2	47.3	9180	97	HSU93237	Human menin	
C 15	1253.2	45.2	203300	85	AC000134	Homo sapi	
C 16	992	35.8	166269	82	AP001187	Homo sapi	
C 17	976.6	35.2	73431	82	AP000575	Homo sapi	
C 18	863.8	31.2	986	9	AR081881	Sequence	
C 19	847	30.6	171980	82	AP000928	Homo sapi	
20	749.6	27.0	2551	8	AF212919	Danio rer	
21	746.4	26.9	812	9	AR081882	Sequence	
22	746.4	26.9	2547	8	DR245952	Danio rer	
23	521.2	18.8	71297	78	AC091082	Homo sapi	
24	506.6	18.3	109290	84	HS838114	Homo sapien	
25	492.2	17.8	538	9	AR081880	Sequence	
26	470	17.0	73431	82	AP000575	Homo sapi	
27	466.2	16.8	513	93	HSA297485	Homo sapi	
28	464	16.7	938	93	HSA297487	Homo sapi	
29	463	16.7	501	93	HSA297486	Homo sapi	
30	463	16.7	816	93	HSA297489	Homo sapi	
31	463	16.7	944	93	HSA297488	Homo sapi	
C 32	453.4	16.4	9286	94	AF109390	Mus muscu	
C 33	453.4	16.4	132365	94	AC006956	Mus muscu	
34	447	16.1	6779	94	AF093756	Mus muscu	
35	446.8	16.1	6736	94	AF024513	Mus muscu	
36	308.4	11.1	4268	4	AB040816	Drosophill	
C 37	231.4	8.3	57803	65	AC018177	Drosophill	
38	231.4	8.3	148781	4	AC008326	Drosophill	
39	231.4	8.3	174287	4	AC007977	Drosophill	
40	231.4	8.3	267518	4	AE003616	Drosophill	
41	214	7.7	157454	82	AP001462	Homo sapi	
42	214	7.7	178100	82	AP001558	Homo sapi	
43	207.4	7.5	209	92	HSA132593	Homo sapi	
44	207.4	7.5	209	92	HSA132594	Homo sapi	
C 45	171	6.2	71297	78	AC091082	Homo sapi	

RESULT 1

HSU93236

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

Location/Qualifiers

1..2772

/organism="Homo sapiens"

/db_xref="taxon:9606"

ALIGNMENTS

HSU93236

Human menin (MEN1) mRNA, complete cds.

2772 bp

mrna

PRI

19-APR-1997

1 (bases 1 to 2772)

Chandrasekharappa,S.C., Guru,S.C., Manickam,P., Olufemi,S.-E., Collins,F.S., Emmert-Buck,M.R., DeBelenko,L.V., Zhuang,Z., Lubensky,I.A., Liotta,L.A., Crabtree,J.S., Wang,Y., Roe,B.A., Weismann,J., Boguski,M.S., Agarwal,S.K., Kester,M., Kim,Y.S., Heppner,C., Dong,Q., Spiegel,A.M., Burns,A.L. and Marx,S.J.

Positional cloning of the gene for multiple endocrine neoplasia-type 1

Science 276 (5311), 404-407 (1997)

97258940

2 (bases 1 to 2772)

Collins,F.S.

Direct Submission

Submitted (13-MAR-1997) National Human Genome Research Institute, Bldg 38A, Room 605, National Institutes of Health, Bethesda, MD 20892, USA

chromosome="11"

/map="11q13"

1..87

/gene="MEN1"

/number=1

1..2772

/gene="MEN1"

88..555

/gene="MEN1"

/number=2

111..1943

/gene="MEN1"

/codon_start=1

/product="menin"

/protein_id="AAC51228.1"

/db_xref="GI:1945387"

/translation="MGLKAAQKTLFPLRSIDDDVRLFAAEELGREGPDVLVLSVLGFLV EHLFANRVIPITNPVLTLPSPADPPGLGTLFPPVADLSIIAALYARTTAQIRGAVD LSLYPREGVSSRELKVKYSDVWVNSLSRSYFKDRAHIQSLFSFTGTKLDSGVAFA VYGAQALGRDVLHIALSDHAWVFGPNGEQTAEVTHGKGNEDRRGOTVNAQVAER SWLYLKGSMRCDRKMEVAFMVCALNPISIDLHTDLSLELLQLOKLLWLLYDHLERY PNALGNADLELEPTGPRDPLTLYHKGIAAKTYRDEHLYPYWLAGYHCRNRV REALQAWADATVIODYNYCRDEDEIYKEFFEVANDVIPNLLKEASLLLEAGEPGE QSQGTQSGSALQDPECFALLRFDYDGCKWEESGPTVLHVGNATFLVQSLGEPGQ VQKRVIRISREAAEAAPAEPEEARERRGRRESKPEEPKPPKPAIDKGLGTGQ GAVSGPPRPKPGTVAQTARGPEGSTAQVPAAPASPPPEGVLTFOSEKMKMGKELLY ATKINSSAIKLQLTAQSQVMKKQKVSTPSTLSPFLKQRKGI."

175

/gene="MEN1"

/standard_name="L22R"

/note="Lys22Arg"

/replace="g"

357..360

/gene="MEN1"

/standard_name="357del4"

/note="4 bp deletion in kindred 0176"

/replace=""

416

/gene="MEN1"

/standard_name="416delC"

/note="1 bp deletion"

/replace=""

464..466

/gene="MEN1"

/standard_name="K119del"

/note="3 bp deletion"

/replace=""

512

/gene="MEN1"

/standard_name="512delC"

/note="1 bp deletion"

/replace=""

556..764

/gene="MEN1"

/number=3

704

/gene="MEN1"

/standard_name="W198X"

/note="Trp198stop"

/replace="a"

735..738

/gene="MEN1"

/standard_name="735del4"

/note="4 bp deletion"

/replace=""

765..893

/gene="MEN1"

/number=4

894..934

/gene="MEN1"

/number=5

935..1022

/gene="MEN1"

exon /number=6
1023..1159
/gene="MEN1"
variation /number=7
1132
/gene="MEN1"
/standard_name="1132delg"
/note="1 bp deletion"
/replace=""
exon 1160..1295
/gene="MEN1"
variation /number=8
1197..1199
/gene="MEN1"
/standard_name="E363del"
/note="3 bp deletion"
/replace=""
exon 1296..1460
/gene="MEN1"
variation /number=9
1416
/gene="MEN1"
/standard_name="W436R"
/note="Trp436Arg"
/replace="c"
variation 1417
/gene="MEN1"
/standard_name="W436X"
/note="Trp436stop"
exon 1461..2772
/gene="MEN1"
variation /number=10
1689
/gene="MEN1"
/standard_name="R527X"
/note="Arg527stop"
/replace="t"
BASE COUNT 577 a 909 c 779 g 507 t
ORIGIN

Query Match 100.0%; Score 2772; DB 97; Length 2772;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2772; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GGTGTCGGAGCGCGGACCTAGACATCCAGAGCCAGCAGCGCGCCCGCC 60
DB 1 GGTGTCGGAGCGCGGACCTAGACATCCAGAGCCAGCAGCGCGCCCGCC 60
QY 61 ACTATTTCCAGGCTCTGCGGGCAGGGCGCGCCACCGCCGCATGGGCTGA 120
DB 61 ACTATTTCCAGGCTCTGCGGGCAGGGCGCGCCACCGCCGCATGGGCTGA 120
QY 121 AGCGCGCCAGAGAGCGTGTTCGCGCTCGCTCCATCGACAGCGTGTGCGCTGTTTG 180
DB 121 AGCGCGCCAGAGAGCGTGTTCGCGCTCGCTCCATCGACAGCGTGTGCGCTGTTTG 180
QY 181 CTGCGGAGCTGGCGGAGAGAGCGGACCTGGTGTCTTCTTCTTGTGCGGCTTCG 240
DB 181 CTGCGGAGCTGGCGGAGAGAGCGGACCTGGTGTCTTCTTCTTGTGCGGCTTCG 240
QY 241 TGGAGCATTTTCTGGCTGTCAACCGCGTTCATCCCTACCAACGTTCCCGAGCTCACCTTCC 300
DB 241 TGGAGCATTTTCTGGCTGTCAACCGCGTTCATCCCTACCAACGTTCCCGAGCTCACCTTCC 300
QY 301 AGCGCGCCCGCCCGCCAGCGCTTGGCGGCTCACCTTCTTCTTGTGCGGCGGCGCTGT 360
DB 301 AGCGCGCCCGCCCGCCAGCGCTTGGCGGCTCACCTTCTTCTTGTGCGGCGGCGCTGT 360
QY 361 CTATCATCGCGCCCTCTATGCGCGCTTACCGCCAGATCCGAGCGCGCTGACCTGT 420
DB 361 CTATCATCGCGCCCTCTATGCGCGCTTACCGCCAGATCCGAGCGCGCTGACCTGT 420

QY 421 CCCTCTATCTCCAGAAGGGGTGTCTCCAGCGGTGAGCTGGTGAAGAAGGTCTCCGATG 480
DB 421 CCCTCTATCTCCAGAAGGGGTGTCTCCAGCGGTGAGCTGGTGAAGAAGGTCTCCGATG 480
QY 481 TCATATGGAACAGCCTCAGCCGCTCTACTTCAAGGATCGGGCCACATCCAGTCCCTCT 540
DB 481 TCATATGGAACAGCCTCAGCCGCTCTACTTCAAGGATCGGGCCACATCCAGTCCCTCT 540
QY 541 TCAGCTTCATCAGCGCACCAAAATTGGACAGCTCCGGTGTGGCTTGTGCTGCTGGGG 600
DB 541 TCAGCTTCATCAGCGCACCAAAATTGGACAGCTCCGGTGTGGCTTGTGCTGCTGGGG 600
QY 601 CCTGCCAGCCCTGGGCTCTCCGGGATGTCCACCTCGCCCTGTCTGAGGATCATGCTGGG 660
DB 601 CCTGCCAGCCCTGGGCTCTCCGGGATGTCCACCTCGCCCTGTCTGAGGATCATGCTGGG 660
QY 661 TAGTGTGGGCCCCAATGGGAGCAGACAGTGAAGTCACTGGGACGCGAAGGCAAG 720
DB 661 TAGTGTGGGCCCCAATGGGAGCAGACAGTGAAGTCACTGGGACGCGAAGGCAAG 720
QY 721 AGSACGCGAGGGCGCAGACAGTCAATGCCGCTGTGGCTGAGCGGAGCTGGCTGTACTCA 780
DB 721 AGSACGCGAGGGCGCAGACAGTCAATGCCGCTGTGGCTGAGCGGAGCTGGCTGTACTCA 780
QY 781 AAGGATCATACATGCGCTGTGACCGCAAGATGGAGTGGCGTTTCATGTTGTGCCATCA 840
DB 781 AAGGATCATACATGCGCTGTGACCGCAAGATGGAGTGGCGTTTCATGTTGTGCCATCA 840
QY 841 ACCCTTCCATTGACCTGTCACACCGACTCGCTGGAGTTCGAGCTGACGAGAACTGC 900
DB 841 ACCCTTCCATTGACCTGTCACACCGACTCGCTGGAGTTCGAGCTGACGAGAACTGC 900
QY 901 TCTGGCTGCTCTATGACCTGGGACATCTCGAAAGGTACCCCATGGCTTTAGGAACTGG 960
DB 901 TCTGGCTGCTCTATGACCTGGGACATCTCGAAAGGTACCCCATGGCTTTAGGAACTGG 960
QY 961 CAGATCTAGAGAGCTGGAGCCACCCCTGGCGGCGCAGACCCACATCCCTCTACCCACA 1020
DB 961 CAGATCTAGAGAGCTGGAGCCACCCCTGGCGGCGCAGACCCACATCCCTCTACCCACA 1020
QY 1021 AGGGCAFTGCCTCAGCCAAAGACTACTATCGGGATGAACACATCTACCCCTACATGTACC 1080
DB 1021 AGGGCAFTGCCTCAGCCAAAGACTACTATCGGGATGAACACATCTACCCCTACATGTACC 1080
QY 1081 TGGCTGGCTTACCACCTGTCCGAACCGCAATGTGGGGAAGCCCTGCGAGGCTGGGCGGACA 1140
DB 1081 TGGCTGGCTTACCACCTGTCCGAACCGCAATGTGGGGAAGCCCTGCGAGGCTGGGCGGACA 1140
QY 1141 CGGCCACTGTCTATCCAGGACTACAACCTACTGCGGGAAGACGAGGAGATCTACAAGGAGT 1200
DB 1141 CGGCCACTGTCTATCCAGGACTACAACCTACTGCGGGAAGACGAGGAGATCTACAAGGAGT 1200
QY 1201 TCTTTGAAGTAGCCAAATGATGTATCCCAACTGCTGTAAGGAGGAGGAGGCTGCTGG 1260
DB 1201 TCTTTGAAGTAGCCAAATGATGTATCCCAACTGCTGTAAGGAGGAGGAGGCTGCTGG 1260
QY 1261 AGCGGGCGCAGAGCGCGGGGAGCAAGCCAGGCGACCCAGAGCCCAAGGTTCCGCC 1320
DB 1261 AGCGGGCGCAGAGCGCGGGGAGCAAGCCAGGCGACCCAGAGCCCAAGGTTCCGCC 1320
QY 1321 TCCAGGACCTGTGAGTCTTTCGCCCACTGCTGCGATTCTTACGACGGCATCTGCAATGGG 1380
DB 1321 TCCAGGACCTGTGAGTCTTTCGCCCACTGCTGCGATTCTTACGACGGCATCTGCAATGGG 1380
QY 1381 AGGAGGCGAGTCCCGACCGCTGTGCTGACGTTGGGCTGGGCGCACTTTCTTGTGCTGCTCC 1440
DB 1381 AGGAGGCGAGTCCCGACCGCTGTGCTGACGTTGGGCTGGGCGCACTTTCTTGTGCTGCTCC 1440
QY 1441 TAGGCGGCTTTGAGGCGACAGGTCCGCGACAGGTGCGCATGTAGTGGCCGAGGCGGAGG 1500
DB 1441 TAGGCGGCTTTGAGGCGACAGGTCCGCGACAGGTGCGCATGTAGTGGCCGAGGCGGAGG 1500

EHFLAVNRVITPNPELTQSPAPDPGGTLTFPVADLSIAALYARFQAIRGAVD
LSLYPREGVSRLVKSVDVWINSLSRSYFKDRHQISLFSFTGTGKLDSSGVAPA
VVGACALGLRDVHLALSDHAWKLYLKGYSMRCDKMEVAFWVCAINPSIDLHTDS
LELLQOKLLWLLDGLHLERYPMALGNLADLELEPTGRPDPLTLHKIGIASAKT
YREBHIYPYMLAGYHCRNRNREALQAWADTIVIODYINCREDEIYKEFFEVA
DVIENLUKEAASLLEAGEERPEQSQSGSALQDPECFALHLLRFYDGIKWEBS
PTVLHVWATFLVQSLGRFGQYKRVIRYREAEAAEAEPEWGEAREGRGPRR
ESKPEEPPLPKPLDGLGTGQAVSGPPRPPTVAGTARGPEGSGTAQVPAPAS
PPPEGPVLTFQSEKMKMKELLVATKINSSAIKLQTLAQSQVQMKKQKRVSPSYTILS
FLKQRKGL"

BASE COUNT 548 a 865 c 729 g 491 t
ORIGIN

Query Match		86.8%;	Score 2406;	DB 91;	Length 2633;
Best Local Similarity		95.7%;	Pred. No. 0;		
Matches 2579;		Conservative 0;	Mismatches 5;	Indels 110;	Gaps 6;
QY	80	GGGCGAGGCGCCGCCACCCGCCCATGGGCTGAAGGCGGCCAGAGACGCT	139		
Db	47	GGCGGAGGCGCCGCCACCCGCCCATGGGCTGAAGGCGGCCAGAGACGCT	106		
QY	140	GTTCGGCTGCCCTCATCGACGAGTGGTGGCGCTGTTTGTGCGGAGCTGGGCGGAGA	199		
Db	107	GTTCGGCTGCCCTCATCGACGAGTGGTGGCGCTGTTTGTGCGGAGCTGGGCGGAGA	166		
QY	200	GGAGCGGACCTGGTGTCTCTTCTTGTGTGGTGTGGTGTGGAGCATTTTCTGCGTGT	259		
Db	167	GGAGCGGACCTGGTGTCTCTTCTTGTGTGGTGTGGTGTGGAGCATTTTCTGCGTGT	226		
QY	260	CAACGCGGTCTATCCCTACCAAGCTTCCGAGCTCACTTTCCAGCCAGCCCGCCCGCA	319		
Db	227	CAACGCGGTCTATCCCTACCAAGCTTCCGAGCTCACTTTCCAGCCAGCCCGCCCGCA	286		
QY	320	CCGCGCTGGCGGCTCACTTCCGCTGGCGGACCTGTCTATCATCGCGCGCTCTA	379		
Db	287	CCGCGCTGGCGGCTCACTTCCGCTGGCGGACCTGTCTATCATCGCGCGCTCTA	346		
QY	380	TGCGCGCTTCAACGCGCCAGATCGAGGCGCGTGCACCTGTCCCTCTATCTCGAGAAGG	439		
Db	347	TGCGCGCTTCAACGCGCCAGATCGAGGCGCGTGCACCTGTCCCTCTATCTCGAGAAGG	406		
QY	440	GGGTGTCTCCAGCGGTGAGCTGGTGAAGAAGTCTCCGATGTCAATGAAACAGCCTCAG	499		
Db	407	GGGTGTCTCCAGCGGTGAGCTGGTGAAGAAGTCTCCGATGTCAATGAAACAGCCTCAG	466		
QY	500	CCGCTCTCTACTTCAAGGATCGGCGCCACATCCAGTCCCTCTTCACTTCACTACAGGCAC	559		
Db	467	CCGCTCTCTACTTCAAGGATCGGCGCCACATCCAGTCCCTCTTCACTTCACTACAGGCAC	526		
QY	560	CAAAATTGGACAGCTCCGCTGTGGCTTTGCTGTGGTGGGCGCTGCCAGGCCCTGGTCT	619		
Db	527	CAAAATTGGACAGCTCCGCTGTGGCTTTGCTGTGGTGGGCGCTGCCAGGCCCTGGTCT	586		
QY	620	CCGGATGTCCACCTCGGCCCTGTGTGAGGATCATGCTGGGTAGTGTGTTGGGCCCAATGG	679		
Db	587	CCGGATGTCCACCTCGGCCCTGTGTGAGGATCATGCTGGGTAGTGTGTTGGGCCCAATGG	624		
QY	680	GGAGCAGACAGCTGAGGTACCTCGCACGGCAAGGCAACGAGGACCGGCGGCGCAGAC	739		
Db	625	-----	624		
QY	740	AGTCAATGCCGTGTGGCTGAGCGGAGCTGGCTGTACCTGAAAGGATCATACATGCGCTG	799		
Db	625	-----GGAGCTGGCTGTACCTGAAAGGATCATACATGCGCTG	661		
QY	800	TGACCGCAAGATGGAGGTGGCGTTGATGTGTGCGCATCAACCCCTTCATTTGACCTGCA	859		
Db	662	TGACCGCAAGATGGAGGTGGCGTTGATGTGTGCGCATCAACCCCTTCATTTGACCTGCA	721		
QY	860	CACCGACTCGCTGGAGCTTCTCAGCTCGACGACGAGCTGCTGGCTGCTCTATGACCT	919		
Db	722	CACCGACTCGCTGGAGCTTCTCAGCTCGACGACGAGCTGCTGGCTGCTCTATGACCT	781		

QY	920	GGGACATCTGGAAGGTACCCCATGSCCTTAGGAACTGGCAGATCTTAGAGGAGCTTGA	979		
Db	782	GGGACATCTGGAAGGTACCCCATGSCCTTAGGAACTGGCAGATCTTAGAGGAGCTTGA	841		
QY	980	GCCCCCCTGGCCGCCAGACCCACTCACCTCTACCAAGAGGCAATTCCTTCAGCCAA	1039		
Db	842	GCCCCCCTGGCCGCCAGACCCACTCACCTCTACCAAGAGGCAATTCCTTCAGCCAA	901		
QY	1040	GACCTACTACTGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCTACCATGTG	1099		
Db	902	GACCTACTACTGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCTACCATGTG	961		
QY	1100	CAACGCAATGTGCGGAAGCCCTGCAGGCTTGGCGGACACGGCCACTGTTCATCCAGGA	1159		
Db	962	CAACGCAATGTGCGGAAGCCCTGCAGGCTTGGCGGACACGGCCACTGTTCATCCAGGA	1021		
QY	1160	CTAACACTACTGCCGGGAAGAGAGAGATCTACAGGAGTTCTTTTGAAGTAGCCAAATGA	1219		
Db	1022	CTAACACTACTGCCGGGAAGAGAGAGATCTACAGGAGTTCTTTTGAAGTAGCCAAATGA	1081		
QY	1220	TGTCATCCCAACCTGCTGAAGGAGCAGCAGCTTGTCTGGAGGCGGCGAGGAGCGGC	1279		
Db	1082	TGTCATCCCAACCTGCTGAAGGAGCAGCAGCTTGTCTGGAGGCGGCGAGGAGCGGC	1141		
QY	1280	GGGGAGCAAAAGCCAGGCGCACAGAGCCCAAGTTCGCGCCCTCCAGGACCCCTGAGTGT	1339		
Db	1142	GGGGAGCAAAAGCCAGGCGCACAGAGCCCAAGTTCGCGCCCTCCAGGACCCCTGAGTGT	1201		
QY	1340	CGCCCACTGTCTGCGATTTCTACGAGGCACTTGCAAATGGAGGAGGAGTCCCAAGCC	1399		
Db	1202	CGCCCACTGTCTGCGATTTCTACGATGGCATCTGCAATGGAGGAGGAGTCCCAAGCC	1261		
QY	1400	TGTGTGTGACGTGGGCTGGGCCACCTTTCTTGTGCACTTCTTAGGCCGTTTGGAGGACA	1459		
Db	1262	TGTGTGTGACGTGGGCTGGGCCACCTTTCTTGTGCACTTCTTAGGCCGTTTGGAGGACA	1321		
QY	1460	GGTGGGCAAGAGGTGCGCATAGTAGCGAGAGAGGCGGCGGCGGCGGCGGAGGAGCC	1519		
Db	1322	GGTGGGCAAGAGGTGCGCATAGTAGCGAGAGAGGCGGCGGCGGCGGAGGAGCC	1381		
QY	1520	GTGGGCGGAGGAAGCCCGGGAAGGCGGCGGCGGCGGCGGCGGCGGAGTCCCAAGCCAGA	1579		
Db	1382	GTGGGCGGAGGAAGCCCGGGAAGGCGGCGGCGGCGGCGGCGGAGTCCCAAGCCAGA	1441		
QY	1580	GGAGCCCGCGCGCCCAAGAACCCAGTGTGACAGGCGCTGGGCGACCGCGGAGGTCG	1639		
Db	1442	GGAGCCCGCGCGCCCAAGAACCCAGTGTGACAGGCGCTGGGCGACCGCGGAGGTCG	1501		
QY	1640	AGTGTGAGGACCCCGCGGAAGCCCTCTGGGACTGTGCTGGCAGACAGCCGAGGCGCTGA	1699		
Db	1502	AGTGTGAGGACCCCGCGGAAGCCCTCTGGGACTGTGCTGGCAGACAGCCGAGGCGCTGA	1561		
QY	1700	AGGTGCGCAGCAGGCTCAGGTGCCAGCACCCCGCAGCATCACACCCTGGAGGAGTCCAGT	1759		
Db	1562	AGGTGCGCAGCAGGCTCAGGTGCCAGCACCCCGCAGCATCACACCCTGGAGGAGTCCAGT	1621		
QY	1760	GCTCACTTTTCCAGAGTGAGAAGATGAAGGCGCATGAAGGAGCTGCTGGTGGCCACCAAGAT	1819		
Db	1622	GCTCACTTTTCCAGAGTGAGAAGATGAAGGCGCATGAAGGAGCTGCTGGTGGCCACCAAGAT	1681		
QY	1820	CAACTCGAGCGCATCAAGCTGCACTACGCGCACAGTGCAGTGCAGATGAAGAAGCA	1879		
Db	1682	CAACTCGAGCGCATCAAGCTGCACTACGCGCACAGTGCAGTGCAGATGAAGAAGCA	1741		
QY	1880	GAAAGTGTCCACCCCTAGTGTACTACTGTCTTTCTTCAAGCGGCGAGGAGGCTTCC	1939		
Db	1742	GAAAGTGTCCACCCCTAGTGTACTACTGTCTTTCTTCAAGCGGCGAGGAGGCTTCC	1801		
QY	1940	CTGAACCTACTGGGAGCTTCGAGCCGCTTGTGGGAGCCAGGCTCCG-CCTTAGTCCCCCA	1998		
Db	1802	CTGAACCTACTGGTGTCTCGGACCGCTTGTGGGAGCCAGGCTCCGCCCTTAGTCCCCCA	1861		


```
QY 1108 ATGTGGGGAAGCCCTGAGGCTGGG---CGGACACGGCCCACTGTCTCCAGGACTACA 1164
Db 1108 ATGTGGGGAAGCCCTGAGGCTGGG---CGGACACGGCCCACTGTCTCCAGGACTACA 1164
QY 1165 ACTACTGCGGGGAAGCAGGAGATCTCAAGAGATTTCTTGAAGTAGCAATGATGCA 1224
Db 1165 ACTACTGCGGGGAAGCAGGAGATCTCAAGAGATTTCTTGAAGTAGCAATGATGCA 1224
QY 1225 TCCCAACCTGCTGAAGAGGAGCAGCTTCTGAGGCGGGGAGGAGGCGCGGGGG 1284
Db 1225 TCCCAACCTGCTGAAGAGGAGCAGCTTCTGAGGCGGGGAGGAGGCGCGGGGG 1284
QY 1285 AGCAAGAGCGGACACCGAGAGCCAAAGTTTCGCGCCTCCAGGACCTGAGTCTTCGCCC 1344
Db 1285 AGCAAGAGCGGACACCGAGAGCCAAAGTTTCGCGCCTCCAGGACCTGAGTCTTCGCCC 1344
QY 1345 ACCTGCTGGATTCTACGACGCGCATCTGCAATGGAGGAGGAGGCGAGTCCAGCCTGTGC 1404
Db 1345 ACCTGCTGGATTCTACGACGCGCATCTGCAATGGAGGAGGAGGCGAGTCCAGCCTGTGC 1404
QY 1405 TGCAGTGGGCTGGGCGCAGCTTCTTGTGTCAGTCCCTAGGCGGTTTGTAGGACAGGTGC 1464
Db 1405 TGCAGTGGGCTGGGCGCAGCTTCTTGTGTCAGTCCCTAGGCGGTTTGTAGGACAGGTGC 1464
QY 1465 GGCAGAGGTGGCAGATAGTGGAGGAGGCGGCGGCGGCGGAGGCGGAGGCGGTGGG 1524
Db 1465 GGCAGAGGTGGCAGATAGTGGAGGAGGCGGCGGCGGCGGAGGCGGAGGCGGTGGG 1524
QY 1525 GCGAGGAAGCGGGAAGCGGCGGCGGCGGCGGCGGCGGCGGAGGAGTCCAGGAGGAGC 1584
Db 1525 GCGAGGAAGCGGGAAGCGGCGGCGGCGGCGGCGGCGGCGGAGGAGTCCAGGAGGAGC 1584
QY 1585 CCGCGCGCCCAAGAGCAGCAGTGGAGGCGCTGGGACCGGCGGCGGAGGTGCTGAGTGT 1644
Db 1585 CCGCGCGCCCAAGAGCAGCAGTGGAGGCGCTGGGACCGGCGGCGGAGGTGCTGAGTGT 1644
QY 1645 CAGGACCCCGCGGAGGCGTGGAGGCGCTGGGACCGGCGGCGGAGGTGCTGAGTGT 1704
Db 1645 CAGGACCCCGCGGAGGCGTGGAGGCGCTGGGACCGGCGGCGGAGGTGCTGAGTGT 1704
QY 1705 GCAGCAGGCTCAGTGCAGGACCGGAGGATCAGGACCGGCGGAGGCGGAGGCTGCTCA 1764
Db 1705 GCAGCAGGCTCAGTGCAGGACCGGAGGATCAGGACCGGCGGAGGCGGAGGCTGCTCA 1764
QY 1765 CTTTCCAGAGTGAAGAGTGAAGGCGATGAAGGCGTGTGCTGCGCAGCAGGAGTCAACT 1824
Db 1765 CTTTCCAGAGTGAAGAGTGAAGGCGATGAAGGCGTGTGCTGCGCAGCAGGAGTCAACT 1824
QY 1825 CGAGCGCCATCAAGTGAAGTGAAGGCGATGAAGGCGTGTGCTGCGCAGCAGGAGTCAACT 1884
Db 1825 CGAGCGCCATCAAGTGAAGTGAAGGCGATGAAGGCGTGTGCTGCGCAGCAGGAGTCAACT 1884
QY 1885 TGTCCACCCCTAGTACACTGCTGCTTCTTCTCAAGGCGGAGGCGGAGGCGCTCTGAA 1944
Db 1885 TGTCCACCCCTAGTACACTGCTGCTTCTTCTCAAGGCGGAGGCGGAGGCGCTCTGAA 1944
QY 1945 CTACTGGGAGCTTCGAGCGCTGTGGGAGCAGGCTCCG-CCTTAGTCCCGCAACTCT 2003
Db 1945 CTACTGGGAGCTTCGAGCGCTGTGGGAGCAGGCTCCG-CCTTAGTCCCGCAACTCT 2003
QY 2004 GAGCCCATGTTTGTGCGCCAGCCCAAGGAGGAGGCGCTCAGCTTACCCAAACCTTAGG 2063
Db 2004 GAGCCCATGTTTGTGCGCCAGCCCAAGGAGGAGGCGCTCAGCTTACCCAAACCTTAGG 2063
QY 2064 TTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGATTTTCTCCAGCTCAGAACCCAGG 2123
Db 2064 TTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGATTTTCTCCAGCTCAGAACCCAGG 2123
QY 2124 GCTCTGCCCGAGTCTAGATATAGTCTCTTCTCCAGATATCCAGCGCGCAATGGA 2183
Db 2124 GCTCTGCCCGAGTCTAGATATAGTCTCTTCTCCAGATATCCAGCGCGCAATGGA 2183
QY 2183 GCTCTGCCCGAGTCTAGATATAGTCTCTTCTCCAGATATCCAGCGCGCAATGGA 1999
Db 2183 GCTCTGCCCGAGTCTAGATATAGTCTCTTCTCCAGATATCCAGCGCGCAATGGA 1999
```

```
QY 2184 AACCTCACGCTGGGCTCTAATTACCAGTCTTTAAAGGCCAGCCCTAGAAACCCAGCT 2243
Db 2184 AACCTCACGCTGGGCTCTAATTACCAGTCTTTAAAGGCCAGCCCTAGAAACCCAGCT 2243
QY 2244 CTTCTCTGG-AACCGCTACCTAGAGCCAGCAACGTTTACTCAGGCTCTCTCCAGCTT 2302
Db 2244 CTTCTCTGG-AACCGCTACCTAGAGCCAGCAACGTTTACTCAGGCTCTCTCCAGCTT 2302
QY 2302 CTTCTCTGG-AACCGCTACCTAGAGCCAGCAACGTTTACTCAGGCTCTCTCCAGCTT 2361
Db 2302 CTTCTCTGG-AACCGCTACCTAGAGCCAGCAACGTTTACTCAGGCTCTCTCCAGCTT 2361
QY 2361 GTAGGAGCTGAGGTTTACCTTAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2178
Db 2361 GTAGGAGCTGAGGTTTACCTTAAACCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2178

RESULT 4
HSURF HSURF 1837 bp mRNA PRI 03-JUL-1998
LOCUS H.sapiens mRNA 3'-region (unknown function).
DEFINITION Y12338
VERSION Y12338.1 GI:2217971
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1837)
AUTHORS Kedra,D.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-1997) D. Kedra, Karolinska Hospital, Department
Of Molecular Medicine, Building L-6, S-171 76 Stockholm, SWEDEN
REFERENCE 2 (bases 1 to 1837)
AUTHORS Kedra,D., Seroussi,E., Fransson,I., Trifunovic,J., Clark,M.,
Lagercrantz,J., Blennow,E., Mehlin,H. and Dumaniski,J.
TITLE The germinal center kinase gene and a novel CDC25-like gene are
located in the vicinity of the PGM gene on 11q13
JOURNAL Hum. Genet. 100 (5-6), 611-619 (1997)
MEDLINE 98001089
FEATURES
source
1..1837
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11q13"
<!.1027
/function="unknown"
/note="unnamed protein product"
/codon_start=2
/protein_id="CAA73007.1"
/db_xref="GI:2217972"
/translation="LGHLEIRYPMALGNLADLEELPTGPRDPFLTLHKGTASAKTYT
RDEHIYPYMLAGYHCRNRNVRREALQAWDATVLDYNYCREDEIEYKFEFVANDY
IPNLLKEASLLLEAGEERPEQSGQSGSALQDPECFHLLRFYDGTCKWEQSPT
PVLHVWATFLVQSLGRPEGOVROKRVIVREAEAEAEPEPMGEARERRRGRRES
KPEPPPPKPKALDKLCTGOGAVSGPRKPGTAVGTARGPEGGTAQVPAPASPP
PEGPVLTFQSEKMKMKELLVATKINSSAIKLQLTAQSOVQMKKQKVSPTSDYTLFL
RQRKRL"
BASE COUNT 412 a 599 c 500 g 326 t
ORIGIN
Query Match 64.2%; Score 1781; DB 97; Length 1837;
Best Local Similarity 99.7%; Pred. No. 0;
Matches 1836; Conservative 0; Mismatches 0; Indels 5; Gaps 5;
QY 917 CTTGGGACATCTGGAAGGTACCCCATGCCCTTAGGAACTGGGAGAGTCTAGAGAGCT 976
Db 917 CTTGGGACATCTGGAAGGTACCCCATGCCCTTAGGAACTGGGAGAGTCTAGAGAGCT 976
QY 977 GGAGCCCACTTGGCGGCGGAGAGCCACTCACCTCTACCACAGGCGATTGCTCCAGC 1036
Db 977 GGAGCCCACTTGGCGGCGGAGAGCCACTCACCTCTACCACAGGCGATTGCTCCAGC 1036
QY 1037 CAAGACCTACTATCGGAGTGAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTG 1096
Db 1037 CAAGACCTACTATCGGAGTGAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTG 1096
QY 121 CAAGACCTACTATCGGAGTGAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTG 180
Db 121 CAAGACCTACTATCGGAGTGAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTG 180
```

1097 TCGCAACCGCAATGTGCGGAAGCCCTGCAGGCTTGGCGGACACAGCCCACTGTCTATCCA 1156
1157 GGACTACAACCTACTGCGCGGAAGACGAGGAGATCTACAAGAGTTCTTTGAAGTAGCCAA 1216
241 GGACTACAACCTACTGCGCGGAAGACGAGGAGATCTACAAGAGTTCTTTGAAGTAGCCAA 300
1217 TGATGTCATCCCAACCTGCTGAAGGAGGAGCAGCCAGCTTCTGAGGCGCGGAGGAGCG 1276
301 TGATGTCATCCCAACCTGCTGAAGGAGGAGCAGCCAGCTTCTGAGGCGCGGAGGAGCG 360
1277 GCGGGGAGCAAAAGCCAGGCGCCACCCAGAGCCAAAGTTTCGCGCCCTCCAGGACCCCTGAGTG 1336
361 GCGGGGAGCAAAAGCCAGGCGCCACCCAGAGCCAAAGTTTCGCGCCCTCCAGGACCCCTGAGTG 420
1337 CTTGCGCCCACTGCTGGGATTTACGACGCGATCTGCAATGGAGGAGGCGAGTCCAC 1396
421 CTTGCGCCCACTGCTGGGATTTACGACGCGATCTGCAATGGAGGAGGCGAGTCCAC 480
1397 GCCTGTCTGCAGTGGGCTGGGCCACCTTTCTGTGCACTCCCTAGGCCGTTTGGAGG 1456
481 GCCTGTCTGCAGTGGGCTGGGCCACCTTTCTGTGCACTCCCTAGGCCGTTTGGAGG 540
1457 ACAGGTGCGGCAGAGGTGGGCATAGTAGCCGAGAGGCGGAGGCGCGGAGGCGCGAGGA 1516
541 ACAGGTGCGGCAGAGGTGGGCATAGTAGCCGAGAGGCGGAGGCGCGGAGGCGCGAGGA 600
1517 GCGTGGGGGAGGAGCCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGAGTCCAGGCC 1576
601 GCGTGGGGGAGGAGCCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGAGTCCAGGCC 660
1577 AGAGGACCCCGCGCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGG 1636
661 AGAGGACCCCGCGCGCGGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGG 720
1637 TGCAGTGTCAAGACCCCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGG 1696
721 TGCAGTGTCAAGACCCCGCGGAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGAGG 780
1697 TGAAGTGGGACGAGCGCTCAGGTGGCGAGCAGCCGCGGAGCATCACCACCGCGGAGGCTC 1756
781 TGAAGTGGGACGAGCGCTCAGGTGGCGAGCAGCCGCGGAGCATCACCACCGCGGAGGCTC 840
1757 AGTGCTCACTTTCCAGAGTGAGAAGATGAAGGCGATGAAGGAGCTGCTGGCGCACCAA 1816
841 AGTGCTCACTTTCCAGAGTGAGAAGATGAAGGCGATGAAGGAGCTGCTGGCGCACCAA 900
1817 GATCAACTCGAGCGCCATCAAGCTGCAACCTCAGGCGACAGTTCGCAAGTGAAGATGAAGAA 1876
901 GATCAACTCGAGCGCCATCAAGCTGCAACCTCAGGCGACAGTTCGCAAGTGAAGATGAAGAA 960
1877 GCAGAAAGTGTCCACCCCTAGTGACTACTGCTCTTCTCAAGCGGAGCGGCGGAGG 1936
961 GCAGAAAGTGTCCACCCCTAGTGACTACTGCTCTTCTCAAGCGGAGCGGCGGAGG 1020
1937 CCTCTGAACCTACTGGGACCTTCGAGCGCTTGTGGGACCGAGGCTCCG-CCTTAGTCCC 1995
1021 CCTCTGAACCTACTGGGACCTTCGAGCGCTTGTGGGACCGAGGCTCCGCGCTTAGTCCC 1080
1996 CCAACTCTGAGCCCATTTCTGCGCCCGAGCCCAAGGGGAGAGCGCTTCACCTTACCCAA 2055
1081 CCAACTCTGAGCCCATTTCTGCGCCCGAGCCCAAGGGGAGAGCGCTTCACCTTACCCAA 1140
2056 ACCCTAGTTCGCGGTCGCGAGTACAGTCTGTATCAAAACCGAGATTTCTCCAGGCTCAG 2115
1141 ACCCTAGTTCGCGGTCGCGAGTACAGTCTGTATCAAAACCGAGATTTCTCCAGGCTCAG 1200
2116 AACCCAGGCTCTGCGCCCGAGTCGTTAGATAATAGGTCTCTTCTCCCAAGATCCAGCGCG 2175
1201 AACCCAGGCTCTGCGCCCGAGTCGTTAGATAATAGGTCTCTTCTCCCAAGATCCAGCGCG 1260

2176 CCAATGGAAACCTCAGCTGGGTCTTAATTAACAGTCTTTTAAAGGCCACGCCCTTAGAAA 2235
1261 CCAATGGAAACCTCAGCTGGGTCTTAATTAACAGTCTTTTAAAGGCCACGCCCTTAGAAA 1320
2236 CCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGCAACGTTTACTCAGGGCTCCCTC 2295
1321 CCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGCAACGTTTACTCAGGGCTCCCTC 1380
2296 CCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGCAACGTTTACTCAGGGCTCCCTC 2355
1381 CCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGCAACGTTTACTCAGGGCTCCCTC 1439
2356 CCGGGAGGCTTAGGACCTCAGCCCTCAGGAGTATATTTCCGCACTTCCAGAAATTTCCAT 2415
1440 CCC-GGAGGCTTAGGACCTCAGCCCTCAGGAGTATATTTCCGCACTTCCAGAAATTTCCAT 1498
2416 ATCTTGGAAATCAAGCTCCTCGCCCAATAACTTCAGTCTCTCTTCCAGAAATTTGGAA 2475
1499 ATCTTGGAAATCAAGCTCCTCGCCCAATAACTTCAGTCTCTCTTCCAGAAATTTGGAA 1557
2476 ATCTTGGAAATCAAGCTCCTCGCCCAATAACTTCAGTCTCTCTTCCAGAAATTTGGAA 2535
1558 ATCTTGGAAATCAAGCTCCTCGCCCAATAACTTCAGTCTCTCTTCCAGAAATTTGGAA 1617
2536 TGAGCATCTCTGAGCCCGCCCTCTTCTGACGAAACTTGGCCCGGATCAGAGCAGGACCT 2595
1618 TGAGCATCTCTGAGCCCGCCCTCTTCTGACGAAACTTGGCCCGGATCAGAGCAGGACCT 1677
2596 CCCTTCGAGCCCTCTGGGAACCTCCAGAGGTCGAGGAGTCCAGGAGCATCCCGGAGGA 2655
1678 CCCTTCGAGCCCTCTGGGAACCTCCAGAGGTCGAGGAGTCCAGGAGCATCCCGGAGGA 1737
2656 AATCTGCAGAGGCTTAGGAGTGGTGACAAAGAGCTGATCTCTCTCTCTTCTTGTACATA 2715
1738 AATCTGCAGA-GGGTAGGAGTGGTGACAAAGAGCTGATCTCTCTCTTGTGTACATA 1796
2716 GATTTATTTTCTAGTTCCTCAAGAAAGATGAATATTTGTT 2756
1797 GATTTATTTTCTAGTTCCTCAAGAAAGATGAATATTTGTT 1837

RESULT 5

AB023400 3067 bp mRNA ROD 02-MAR-2000
LOCUS Rattus norvegicus MEN1 mRNA for menin, complete cds.
DEFINITION
ACCESSION AB023400
VERSION AB023400.1 GI:5360728
KEYWORDS MEN1; menin.
SOURCE Rattus norvegicus (strain:Wistar) Brain cDNA to mRNA.
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
Rattus.
REFERENCE 1 (sites)
AUTHORS Maruyama,K., Tsukada,T., Hosono,T., Ohkura,N., Kishi,M., Honda,M.,
Nara-Ashizawa,N., Nagasaki,K. and Yamaguchi,K.
TITLE Structure and distribution of rat menin mRNA
JOURNAL Mol. Cell. Endocrinol. 156 (1-2), 25-33 (1999)
MEDLINE 20077667
REFERENCE 2 (bases 1 to 3067)
AUTHORS Maruyama,K., Tsukada,T., Nagasaki,K., Maruyama,K., Tsukada,T.,
Nagasaki,K., Ohkura,N., Honda,M., Nara,N., Hosono,T., Kishi,M.,
Kazuki,S., Futami,H. and Yamaguchi,K.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-1999) to the DDBJ/EMBL/GenBank databases. Kouji
Maruyama, National Cancer Center Research Institute, Growth Factor
Division; 5-1-1 Tsukiji, Chuo-ku, Tokyo 104-0045, Japan
(E-mail:kmaruyamegan2.res.ncc.go.jp, tel:81-3-3542-2511(ex.4302),
Fax:81-3-3542-8170)
FEATURES
Location/Qualifiers
1..3067
/organism="Rattus norvegicus"
/strain="Wistar"

gene	/db_xref="taxon:10116" /tissue_type="Brain" 470..2302 /gene="MEN1" 470..2302 /gene="MEN1" /codon_start=1 /product="menin" /protein_id="BAA82134.1" /db_xref="GI:5360729"	
CDS	/translation="MGLKAAOKTLPLRSIDDVVRFLFAAELGREPDLVLSLVGFV EHLAVNRVPTNPVELTQFSPAPDPGGLTYFPVADLSIAALYARFTAOIRGAVD LSLPREGGVSSRELVKKYSDDVTWNSLSRSYFKDRAHIQSLFSFTGTGKLDSSGVAFV VVGCAQLGRDVLHLSAISEDHAWVFGNSGEQTAETVWHGKGNDRGQTVNAGVAER SWLYLKSGMYRCDRKMVAFVCAINPISDILHSDLSLELLQLQKLLWLYDLGLYERY PMALGNLADLELEPTGPRDPLUTLYLHKGIASAKTYQDEHYIPWYLAGYHGRNRV REALQWADTATYIQDYNCREDEIYKEFFEVANDVTPNLLKEAASLLAEAGERPE QAQGTGGQALQDQFAHLIRFYDGIKWEESPTPLVHGWNATFLVQSLGRFEGQ VRQKVHIVSRSEAEEAEPEWGEARERRGRPRRESKPEEPPPPKPKPALDKFGSGQ SAGSGPRKTSGTGVTARGTEVSSAAQAPAPAAPPPPEPVLTFQSEKMKMKKELLY ATKINSAILQLTAOSQVMKKQKVSPTSDYTLFLKLRQKGL"	
BASE COUNT	672 a	904 c 881 g 510 t
ORIGIN		
Query Match	59.5%;	Score 1648.6; DB 94; Length 3067;
Best Local Similarity	78.4%;	Pred. No. 0;
Matches 2141; Conservative	0;	Mismatches 519; Indels 71; Gaps 11;
QY	43	GCAGCGCGCCGCCGCACTATTTCAGGCTCTGCGGGGAGGGCGCGCCACCGC 102
Db	402	GGAGGGGGCTCTCGCGGAGCGCCGACCCCTCCCTCCCTGCGCTGGCCGACCGC 461
QY	103	CCGCGCCCATGGGCTGAAGCGCCGCCAGACGCTGTTCGCCGCTGCGCTCCATCGACG 162
Db	462	CCGCGCGCATGGGCTGAAGCGCGCCGAGACGCTGTTCCTCTGCGCTCTATCGACG 521
QY	163	ACGTGTGCGGCTGTGTGTCGCCAGCTGGCGCGAGAGGAGCGGACCTGTGCTCTCTT 222
Db	522	ACGTGTGCGGCTGTGTGTCGCCAGCTGGCGCGAGAGGAGCTGACCTGTGCTCTCT 581
QY	223	CCTTGTGCTGGGCTTGTGAGACATTTTCTGGCTGTCAACCGGCTATCCCTACCAACG 282
Db	582	CCTTGTGCTGGGCTTGTGAGACATTTTCTGGCTGTCAACCGTGTATCCCGACCAACG 641
QY	283	TTCCCGAGCTACCTTCCAGCCAGCCCGCCCGCCGACCTGCGGCGCTCACCTACT 342
Db	642	TGCGGAGCTCACCTTCCAGCCAGCCCGCCGACCTGCTGCTGCTCACCTACT 701
QY	343	TTCCCGTGGCGACCTGTCTATCATCGCGCGCTCTATGCCCGCTTACCGCCGAGATCC 402
Db	702	TCCAGTGGCGGACCTATCCATCATGTGTCGCTCTATGCGCGGATTCACCGCTCAGATCC 761
QY	403	GAGGCGCGCTGCACCTGTCCCTCTATCTCTCAGAAAGGGGTGTCTCCAGCCGTGAGCTGG 462
Db	762	GCGGTGCTGTGACCTCTCCCTCTATCTCTCAGAGGGGGGGTTCACAGTCCGAACTGG 821
QY	463	TGAAGAGGCTCCGATGTTCATATGGAACAGCTCAGCGCGCTCTACTTCAAGGATCGGG 522
Db	822	TAAAAAGGCTCGGATGTTCATATGGAACAGCTCAGCGCGCTCTACTTCAAGGACCGGG 881
QY	523	CCACATCCAGTCCCTCTTACGCTTCATCACAGGACCAAAATTTGGACAGCTCCGGTGTGG 582
Db	882	CCACATCCAGTCTCTCTTACGCTTCATCACAGGACCTAAACTGGACATTCGGCGCGTGG 941
QY	583	CCTTGTGCTGGTGGGCGCTCCAGGCGCTTGGGTCTCCGGGATGTCCACCTCGCCCTGT 642
Db	942	CCTTGTGAGTGGTAGGGCTCCAGGCTCTAGGTCTCAGAGATGTCATCTGCGCCCTGT 1001
QY	643	CTGAGGATCATGCTGGGTAGTGTGTTGGGCCCAATGGGGAGCAGACAGCTGAGGTCACT 702
Db	1002	CCGAGGATCATGCTGGGTAGTGTGTTGGGTCCAAATGGGGAACAGACGCTCAGGTGACCT 1061

QY	703	GGCAGCGCAAGGGCAACGAGGACCGGCGAGAGTCAATGCGGTGTGCTGTGAGC 762
Db	1062	GGCATGGCAAGGCAACGAGGACCGGCGAGAGCCAGACAGTCAACGCTGTGTGGCTGAGC 1121
QY	763	GGAGCTGGCTGTACCTTGAAAGGATCATATGCGCTGTGACCCCAAGATGGAGGTGGCGT 822
Db	1122	GGAGCTGGCTGTACCTTGAAAGGATCGTACATGCGCTGTGACCCCAAGATGGAGGTGGCGT 1181
QY	823	TCATGGTGTGTGCCATCAACCCCTTCCATTGACCTGCACACCGACTCGCTGCGAGCTTCTGC 882
Db	1182	TCATGGTGTGTGCCATCAACCCCTTCCATTGATGCTTTCACACTGATCTTTTGGAACTGTTC 1241
QY	883	AGCTGCAGCAGAGCTGCTCTGCTGCTCTATGACCTGGGACATCTCGAAAGGTACCCCA 942
Db	1242	AACTACAGCAGAGCTGCTCTGCTGCTGTATGACCTAGGACATCTGGAAGGTACCCCA 1301
QY	943	TGGCTTTAGGAACTTGGCAGATCTAGAGAGCTTGAAGAGCTTGGAGCCCACTTGGCGGCGAGACC 1002
Db	1302	TGGCGCTAGGAACTTGGCAGACTTGGAGAGCTTGAAGAGCTTGGAGCCCACTTGGCTGGCGGAGACC 1361
QY	1003	CACCTACCCCTTACCAACAGGGCATTTGCCCTCAGCCAGAGACTTACTATCGGATGAACACA 1062
Db	1362	CGCTACCCCTTATCAAGAGGATTTGCCCTCAGCTAAGACTTACTACAGGATGAACACA 1421
QY	1063	TCTACCCCTTACATGTACCTGGCTGCTACCACTGTCCCAACCGCAATGTGCGGGAAGCCC 1122
Db	1422	TCTACCCCTTACATGTACCTGGCTGCTGATACCATTTGTCGAACCGCAATGTGCTGGAAGCCC 1481
QY	1123	TGAGGCTTGGGGGAGACGGCCACTGTCTATCCAGGACTACAACTACTTGCCTGGGAGAGC 1182
Db	1482	TGAGGCTTGGGGGAGACGGCCACTGTCTATCCAGACTACAACTACTTGCCTGGGAGAGC 1541
QY	1183	AGGAGATCTACAGGAGTCTTTTGAAGTAGCCCAATGATCTATCCCACTCTCTGAAG 1242
Db	1542	AGGAGATCTACAGGAGTCTTTTGAAGTAGCCCAATGATCTATCCCACTCTCTGAAG 1601
QY	1243	AGGAGCTTGTCTGAGGGCGGAGAGCGCGGGGGAGAGCAAGCAAGCGGAGCC 1302
Db	1602	AGGAGCTTGTCTGAGGGCGGAGAGCGCGGGGGAGAGCAAGCAAGCGGAGCC 1661
QY	1303	AGAGCCAAAGTTTCCGCCCTTCCAGGACCTTGTGCTTTCGCCCACTTGTCTGCGATTTCTAGC 1362
Db	1662	AGAGCCAAAGTTTCTGCCCTTCCAGGACCTTGTGCTTTCGCCCACTTGTCTGCGATTTCTAGC 1721
QY	1363	ACGGATCTGCAAAATGGAGGAGGCGAGTCCCACTGCTGTCTGCTGACGCTGGGCTGGGCA 1422
Db	1722	ATGGATCTGCAAAATGGAGGAGGCTAGCCCACTGCTGCTGCTGACGCTGGGCTGGGCA 1781
QY	1423	CCTTCTTGTGAGTCCCTTAGGCGCTTTTGAAGGACAGGTGCGGCGAGAGGTGCGCATAG 1482
Db	1782	CCTTCTTGTGAGTCCCTTAGGCGCTTTGAGGACAGGTGCGGCGAGAGGTGCGCATAG 1841
QY	1483	TGAGCGGAGGCGCGGCGGCGAGGCGGAGGCGGCTGGGGGAGAGCAAGCGGGAAG 1542
Db	1842	TTAGCGGGAAGCGGAGGCGAGAGGCGGGAAGCAATCGGGGGAGCAAGCGGGAAG 1901
QY	1543	GCGGGCGGCGGCGGCGGCGGAGTCCCAAGCCAGGAGGCGGCGGCGGCGGCGGCAAGAC 1602
Db	1902	GTCGTGCGGCGGCTGCCCGCGAGAGTCCCAAGCTTGAAGCTTGAAGGACCACTACCAAGAAC 1961
QY	1603	CAGCACTGACAAAGGCGCTGGGCGCGGCGAGGTGCTGCTGAGGACCGGCGGGAAGC 1662
Db	1962	CTGCGTTGCAAGGCGGCGGCGGCTGAGGCAAGGTGAGGGTGGGAGCTCTTAGGAAA 2021
QY	1663	CTTCTGGGACTGTCTGTCGACAGCGCCGAGGCGCTTGAAGGTGGCAGCAGGCTCAGGTGC 1722
Db	2022	CGTCAAGGACTGTCTCAGGTACTTGCCTGGCACAGAACTGAGTGTCTGCTCAGGCTC 2081
QY	1723	CAGCAGCCGCGGAGTCCACCGCGGAGGCTCAGTGTCTACTTCTTCCAGAGTGAAGAG 1782
Db	2082	CAGCAGCTGAGCAGTCCGCGGAGGCGGCGGCTGCTGCTACTTCTTCCAGAGTGAAGAG 2141
QY	1783	TGAAGGGCATGAAGGAGCTGCTGTTGGGCCCAAGATCAACTCGAGCGCCATCAAGCTGC 1842

Db	130	GCCAAATGATGTATATCCCAACCTGCTGAAGAGGACAGCCAGTTCGTGAGGGGGCGAG	189
Qy	1272	GAGCGCGCGGGGAGCAAAAGCCAGGAGCACCCAGAGCCAAAGTTCGCGCCTCCAGGACCCCT	1331
Db	190	GAGCGCGCGGGGAGCAAAAGCCAGGAGCACCCAGAGCCAAAGTTCGCGCCTCCAGGACCCCT	249
Qy	1332	GAGTGTTCGCCCACTGCTGTGGATTTACGACGGCATCTGCAAAATGGGAGGAGGCGAGT	1391
Db	250	GAGTGTTCGCCCACTGCTGTGGATTTACGACGGCATCTGCAAAATGGGAGGAGGCGAGT	309
Qy	1392	CCACAGCCTGTGTCACGTGGGCTGGGCCACCTTCTTGTGTCAGTCCCTAGGCCGTTT	1451
Db	310	CCACAGCCTGTGTCACGTGGGCTGGGCCACCTTCTTGTGTCAGTCCCTAGGCCGTTT	369
Qy	1452	GAGGACAGGTGCGGCAGAAAGTTCGCGCATAGTGAGCCCGAGAGGCCGAGGCGCGGAGGCC	1511
Db	370	GAGGACAGGTGCGGCAGAAAGTTCGCGCATAGTGAGCCCGAGAGGCCGAGGCGCGGAGGCC	429
Qy	1512	GAGAGCGGTGGGGCGAGGAAGCCCGGGAAGCGCGGCGGGGCCACACGCGCGGAGTCC	1571
Db	430	GAGAGCGGTGGGGCGAGGAAGCCCGGGAAGCGCGGCGGGGCCACACGCGGAGTCC	489
Qy	1572	AAGCCAGAGAGCCCCCGCGGCCAAGAACCCAGCACATGACAAAGGGCTTGGGACCGGGC	1631
Db	490	AAGCCAGAGAGCCCCCGCGGCCAAGAACCCAGCACATGACAAAGGGCTTGGGACCGGGC	549
Qy	1632	CAGGGTCACTGTACGAGACCCCGGGAAGCTCCTGGGACTGTGCTGGCACAGCCCGA	1691
Db	550	CAGGGTCACTGTACGAGACCCCGGGAAGCTCCTGGGACTGTGCTGGCACAGCCCGA	609
Qy	1692	GGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCGAGCATCACACCGCCGAG	1751
Db	610	GGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCGAGCATCACACCGCCGAG	669
Qy	1752	GGTCCAGTGTCACTTTCAGAGTGAGAAGATGAAGGCGATGAAGGAGCTGCTGTGGCC	1811
Db	670	GGTCCAGTGTCACTTTCAGAGTGAGAAGATGAAGGCGATGAAGGAGCTGCTGTGGCC	729
Qy	1812	ACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGGCACACAGTGCAGTCCAGATG	1871
Db	730	ACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGGCACACAGTGCAGTCCAGATG	789
Qy	1872	AAGAAGCAGAAAGTGCACCCCTAGTCACTACACTGTCTTTCCCTCAAGCGCAGCGC	1931
Db	790	AAGAAGCAGAAAGTGCACCCCTAGTCACTACACTGTCTTTCCCTCAAGCGCAGCGC	849
Qy	1932	AAAGGCTCTGAACATACTGGGACATTCGGACCGCTTGTGGGACCCAGGCTTCG - CTTA	1990
Db	850	AAAGGCTCTGAACATACTGGGACATTCGGACCGCTTGTGGGACCCAGGCTTCG - CTTA	909
Qy	1991	GTCCCCCAACTCTGAGCCCATGTCGCCCCAGCCCAAGAGGACAGGCTCACCTCTA	2050
Db	910	GTCCCCCAACTCTGAGCCCATGTCGCCCCAGCCCAAGAGGACAGGCTCACCTCTA	969
Qy	2051	CCCAAAACCTTAGTTCGCCGTCCGAGTACGTCGTATCAAAACCCACGATTTCTCCAG	2110
Db	970	CCCAAAACCTTAGTTCGCCGTCCGAGTACGTCGTATCAAAACCCACGATTTCTCCAG	1029
Qy	2111	CTCAGAACCCAGGGCTCTGCCCGAGTCGTTAGAATATAGTCTCTTCTCCCAAGATCCCA	2170
Db	1030	CTCAGAACCCAGGGCTCTGCCCGAGTCGTTAGAATATAGTCTCTTCTCCCAAGATCCCA	1089
Qy	2171	GCGGGCAATGGAACCTCAGCTGGGTCCCTAATTTACAGTCTTTAAAGGCCAGGCCCT	2230
Db	1090	GCGGGCAATGGAACCTCAGCTGGGTCCCTAATTTACAGTCTTTAAAGGCCAGGCCCT	1149
Qy	2231	AGAAACCCAAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGCTTACTCAGGC	2290
Db	1150	AGAAACCCAAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGCTTACTCAGGC	1209
Qy	2291	TCCTCCAGCTTGTAGGAGCTGAGGTTTACCTTTAAACCCAAAGGGAGCAGAGTCCACCC	2350

Db	1210	TCCTCCAGCTGTGTAGAGCGTGGGTTTTCACCTTTAAACCCAA -GGAGCACAGGTCGCCACC	1262
Qy	2351	TCCAGCCGGGGAGCCTTAGGACACATCAGCCCTCCTAGGAGTATATTTCCGCACCTTCAGAAT	2410
Db	1269	TCCAGCCC -GGGAGCCTTAGGACACATCAGCCCTCCTAGGAGTATATTTCCGCACCTTCAGAAT	1327
Qy	2411	TCCATATCTTTGCGAATCCAAGCTCCCTGCCCCCAATAACTTCAGTCTGCTTCCAGAATT	2470
Db	1328	TCCATATCTTTGCGAATCCAAGCTCCCTGCCCCCAATAACTTCAGTCTGCTTCCAGAATT	1386
Qy	2471	TGGAATCCTAGTCTTCTCTCTCTATCCGAGTCTGGGACACAAAACTTCGCCGCCCA	2530
Db	1387	TGGAATCCTAGTCTTCTCTCTCTATCCGAGTCTGGGACACAAAACTTCGCCGCCCA	1446
Qy	2531	GCCTATCAGCATCCTGAGCCCCGCCCTCTTCTCTGACGAAACTGGCCCGGATCAGAGCAG	2590
Db	1447	GCCTATCAGCATCCTGAGCCCCGCCCTCTTCTGACGAAACTGGCCCGGATCAGAGCAG	1506
Qy	2591	GACCTCCCTTCGAGCCCTCTGGGAACCTCCCGAGAGGTCCAGCCCCATCTCGGAGCATCCCG	2650
Db	1507	GACCTCCCTTCGAGCCCTCTGGGAACCTCCCGAGAGGTCCAGCCCCATCTCGGAGCATCCCG	1566
Qy	2651	GAGGAATCTGCAGAGGGGTAGGAGTGGGTGACAAAGAGCCTGATCTCTCTCTTTGT	2710
Db	1567	GAGGAATCTGCAGA -GGGTAGGAGTGGGTGACAAAGAGCCTGATCTCTCTCTTTGT	1625
Qy	2711	ACATGAGATTTATTTTTCAGTTCACGAAAGATGAATACATTTTGTAAATAAAAAA	2770
Db	1626	ACATGAGATTTATTTTTCAGTTCACGAAAGATGAATACATTTTGTAAATAAAAAA	1685
Qy	2771	AA 2772	
Db	1686	AA 1687	
RESULT	7		
AF109389			
LOCUS	AF109389	2984 bp	08-JUN-1999
DEFINITION	Mus musculus menin (Men1)	mRNA	
ACCESSION	AF109389	mRNA	complete cds.
VERSION	AF109389.1	GI:5006429	
KEYWORDS			
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE			
AUTHORS	Guru,S.C., Crabtree,J.S., Brown,K.D., Dunn,K.J., Manickam,P., Prasad,N.B., Wangsa,D., Burns,A.L., Spiegel,A.M., Marx,S.J., Pavan,W.J., Collins,F.S. and Chandrasekharappa,S.C.		
TITLE	Isolation, genomic organization, and expression analysis of Men1 the murine homolog of the MEN1 gene		
JOURNAL	Mamm. Genome 10 (6), 592-596 (1999)		
MEDLINE	99272818		
PUBMED	10341092		
REFERENCE	2 (bases 1 to 2984)		
AUTHORS	Guru,S.C., Crabtree,J.S., Brown,K.D., Manickam,P., Dunn,K.J., Pavan,W.J., Wangsa,D., Prasad,N., Burns,A.L., Marx,S.J., Spiegel,A.M., Collins,F.S. and Chandrasekharappa,S.C.		
TITLE	Submitted (25-NOV-1998) Genetics and Molecular Biology Branch, National Human Genome Research Institute, NIH, Bldg. 49, Rm. 3C, 9000 Rockville Pike, Bethesda, MD 20892, USA		
JOURNAL	Direct Submission		
FEATURES	Location/Qualifiers		
source	1. .2984		
	/organism="Mus musculus"		
	/strain="129Sv/cj7"		
	/db_xref="taxon:10090"		
	/chromosome="19"		
gene	1. .2984		
	/gene="Men1"		
CDS	422..2257		
	/gene="Men1"		

/note="similar to Homo sapiens MENIN"

/codon_start=1

/product="menin"

/protein_id="A006430"

/db_xref="GI:5006430"

/translation="MGKRAQKTLFPLRSIDVVRFAELREBPDVLVLLSLVLFV
EHFLAVRNIPVNTPELTPSPAPDPGLTYFPVADLSIIAALYARETAQIRGAVD
LSLYPREGVSSRLVKYSDVITWNSLSRSYFKDRAHIQSLFSFITGKLDSSGVAFA
VVGACALGLRDVHLALSDHAWFVNGEOTAEVTHWKGKNGRRGOTVNAVGAER
SWLYLAKSYMCRDKMEVAFMCAINPSIDLTSLLELLOKLLWLLVLDLHLERY
PHALGNLADLEELPPGPRDPLTLVHKGIASAKTYQDEHYIPYWLGYHCRNRV
REALQAWDTAVIDQNTCREDEEYKEFFEVANDVIINLKEASLLEETGEERGE
QAQGTGGSAIDPECFALHRLDYDICKWEEGSPVPLHVGNATFLVQSLGFEQG
VRQKHIVIREAAAEPEFWDARERRRPRRESKPEEPKPPKALDRKPGSGQ
SAGSGPPRTSGTVPTGTGGQEVGNAOAPAPASPPPEPGLTFQSEKMGKMKELL
VATKINSSAIKILQTAQSOVMKKOKVSPSPDYTLFLKLRQKGL"

BASE COUNT 656 a 889 c 859 g 580 t

ORIGIN

Query Match 58.3%; Score 1616.4; DB 94; Length 2984;
Best Local Similarity 79.1%; Pred. No. 4.8e-294;
Matches 2130; Conservative 0; Mismatches 446; Indels 116; Gaps 13;

QY 88 GCCGCCGCCACCGCCCGCCGATGGGCTGAAGCGCCGCCAGAGACGCTGTTCGCCG 147
DB 399 GCCCGCAGCCACCGCCCGCCGATGGGCTGAAGCGCCGCCAGAGACGCTGTTCGCC 458
QY 148 TGGCCTCCATCGAGAGCTGGTGGCTGTTGCTGCGAGCTGGCGCGAGAGAGCGG 207
DB 459 TGGCCTCTATCGAGAGCTGGTGGCTGTTGCTGCGAGAGCTGGCGCGAGAGAGCGG 518
QY 208 ACCTGTGCTCCTTTCCTTGGTGGCTGGCTGGAGCATTTTCTGGCTGTCAACCGCG 267
DB 519 ACCTGTGCTCCTTGGCTGGCTGGCTGGAGCATTTTCTGGCTGTCAACCGCGT 578
QY 268 TCATCCCTACACAGTTCCGAGCTACCTTCAGGCCCGCCCGCCCGCCGAGCGCGCTG 327
DB 579 TCATCCCGCCACAGTCCGAGCTACCTTCAGGCCCGCCCGCCCGCCGAGCGCGCTG 638
QY 328 GCGGCTCACTACTTTCGCTGGCGACCTGTCTATCATCGCGCGCTCTATGCGCGCT 387
DB 639 GTGGCTCACTACTTCCGCTGGCGACCTATTCATCATTTGCTGCTCTATGCGCGCT 698
QY 388 TCACCGCCAGATCCGAGCGCGCTGACCTGTCCCTCTATCTTCGAGAGGGGCTGTCT 447
DB 699 TCACCGCTCAGATCCGCGCGCTGTGACCTCTCCCTCTATCTTCGAGAGGGGCTGT 758
QY 448 CCAGCTGAGCTGGTGAAGAAGTCTCCGATGTCATATGAACAGCTCAGCGGCTCT 507
DB 759 CTAGTCGGAACCTGGTAAAAAGGCTCTCGATGTGTCATATGAACAGCTCAGCGCTCT 818
QY 508 ACTTCAAGGATCGGGCCACATCCAGTCCCTCTTCAGCTTCATCAGAGGACCAATTTG 567
DB 819 ACTTCAAGGACCGGGCCACATCCAGTCCCTCTTCAGCTTCATCAGAGGACCAATTTG 878
QY 568 ACAGCTCGGCTGTGGCTTGTGCTGTGGGCTGCCAGGCGCTGGGCTCCGGGATG 627
DB 879 ACAGCTCGGCGTGGCTTGTGCTGTGGGCTGCCAGGCTCTGGGCTCTCAGAGATG 938
QY 628 TCCACCTCGGCTGTGAGATCATGCTGGGTAGTGTGGTGGCCCAATGGGGAGCAGA 687
DB 939 TCCATCTGGGCTGTGTAAGATCATGCTGGGTGGTGTGGTGGCCCAAGGGGAGCAGA 998
QY 688 CAGCTGAGGTCACTTGGCAGCGGCAAGGCAACAGAGCGCGAGGCGCCAGAGTCAATG 747
DB 999 CAGCTGAGGTGACGTGGCAGCGGCAAGGCAACAGAGCGCGAGGCGCCAGAGTCAATG 1058
QY 748 CCGGTGTGGCTGAGCGAGCTGGCTGTACCTGAAAGGATCATACATCGCTGTGACCGCA 807
DB 1059 CCGGTGTGGCTGAGCGAGCTGGCTGTACCTGAAAGGCTGTACATCGCTGTGACCGCA 1118
QY 808 AGATGGAGGTGGCTTTCATGGTGTGCCATCAACCCCTTCATTTGACCTGCACACCGACT 867

DB 1119 AGATGGAGGTGGGCTTTCATGGTGTGCCATCAACCCCTTCATCGATCTTTCACACTGACT 1178
QY 868 CGCTGGAGCTTTCGACAGCTCAGCAGAGCTGCTGCTGGCTCTATGACCTGGGACATC 927
DB 1179 CTTTGGAACTGTTTGAACACTACGAGAGCTGCTGCTGGCTGTATGACCTGGGACATC 1238
QY 928 TGGAAAGGTACCCCATGGCTTAGGAACTTGGCAGACTTAGAGGAGCTGGAGCCACCC 987
DB 1239 TGGAAAGATACCCCATGGCTTAGGAACTTGGCAGACTTGGAGGAGCTGGAGCTTACC 1298
QY 988 CTGCGCGCCAGACCCACTCACCTCTACCAAGGCGCATTTGCCTCAGCCAAGACTTACT 1047
DB 1299 CGGCGCGCCAGACCCACTCACCTTTATCACAGGGAATTTGCCCTCAGCTAAGACCTTACT 1358
QY 1048 ATCGGATGAACACATCTACCCCTACATGTACCTGGCTGCTACCACTGTGCAACCGCA 1107
DB 1359 ACCAGGATGAACACATCTACCCCTACATGTACCTGGCTGCTACCACTGTGCAACCGCA 1418
QY 1108 ATGTGCGGGAAGCCCTGAGGCTGGCGGACAGCGGCACTGTCTATCCAGGACTTACAAC 1167
DB 1419 ATGTGCGGAAGCCCTGAGGCTGGCGGACAGCGGCACTGTCTATCCAGACTTACAAC 1478
QY 1168 ACTGCGGGAAGAGGAGGATCTACAGGAGTTCTTTGAAGTAGCCAATGATGTCTATCC 1227
DB 1479 ACTGCGGGAAGGAGGATCTACAGGAGTTCTTTGAAGTAGCCAATGATGTCTATCC 1538
QY 1228 CCAACCTGTCTGAAGGAGCGAGCTTGTGGAGGCGGCGAGGCGCGCGGGGAGC 1287
DB 1539 CCAACCTGTCTGAAGGAGCGGCGGAGCTGTAGAACAGGCGGAGGAGCTGGGGAGC 1598
QY 1288 AAGGCCAGGCAAGGCTTCCGCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1347
DB 1599 AAGGCCAGGCAAGGCTTCCGCTTCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1658
QY 1348 TGCTGCGATTTCTAGGAGGAGCTGTGCAAAATGGGAGGAGGAGGAGGAGGAGGAGGAG 1407
DB 1659 TGCTGCGATTTCTAGGAGGAGCTGTGCAAAATGGGAGGAGGAGGAGGAGGAGGAGGAG 1718
QY 1408 ACCTGGGCTGGGCAAGCTTCTTGTGAGTCCCTTAGGCGGCTTTTGAGGAGGAGGAGGAG 1467
DB 1719 ACCTGGGCTGGGCAAGCTTCTTGTGAGTCCCTTAGGCGGCTTTGAGGAGGAGGAGGAG 1778
QY 1468 AGAAGTGGCGATAGTGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1527
DB 1779 AGAAGTGGCGATAGTGAAGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1838
QY 1528 AGAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1587
DB 1839 ATGAAGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1898
QY 1588 CGCGCGCCAAAGAGCCAGCACTGGACAAAGGCGCTGGGCGACCGCGGCGGCGGCTGAGTGT 1647
DB 1899 CACCAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1958
QY 1648 GACCGCGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1704
DB 1959 GACCACTAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2018
QY 1705 GCAGCAGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1764
DB 2019 GCAATGCTGTGAGGCTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2078
QY 1765 CTTTCCAGAGTGAAGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1824
DB 2079 CTTTCCAGAGTGAAGATGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2138
QY 1825 CGAGCGCCATCAAGCTCAACTCACGCGAGCTGCGAAGTGCAGATGAAGAGCAGAAAG 1884
DB 2139 CGAGCGCCATCAAGCTCAACTCACGCGAGCTGCGAAGTGCAGATGAAGAGCAGAAAG 2198
QY 1885 TGTCACCGCTTGTGAGTACACTGTCTTCTCAAGCGGCGAGGAGGAGGAGGAGGAGGAGGAG 1944

Db 2199 TGTCCACACCCAGCGACTACACACTCTCTTTCTTAAGCGGAGCGCAAGGCCCTCTGA - 2257
 QY 1945 CTACTGGGACTTCGGACCGCTTGTGGGACCCAGGCTCCGCTTAGTCCGCCCAACTCTG 2004
 Db 2258 -----ACTCTGATTCGTGGCAAGCCCGAGCCCTCCTCTACTCCAGCAGGCTT 2308
 QY 2005 AGCCCATGTTCTGCCCCAGCCCAAGGGGAGAGGCTTACCTCTACCCAAACCCCTAGT 2064
 Db 2309 CACCTCATCTCTTAATCAAGCCTA-GACTCCGCCCTCAGTCTAGGGAGCCAGAGCT 2367
 QY 2065 TCCGGTCCGAGTACAGTCTGTATCAACACCACAGATTTTCTCCAGCTCAGAACCCAGG 2124
 Db 2368 GTCCGCACTGACTA-----GGAGCCTTCAATTTAGCCCTCTTCAGGCCCCAGAG 2417
 QY 2125 CTCTGCCCGAGTCGTAGATATAGTCTCTTCTCCAGAAATCCAGCGCGCAATGGAA 2184
 Db 2418 CCTGGCTCAGTCTTAGAATGTAGGTGTATCCCG-----ACAGCCCGCAATTTGA 2471
 QY 2185 ACCTCAGCTGGTCTTAATACAGTCTTTAAAGGCCCGAGCCCTTAGAAACCCCAAGCTC 2244
 Db 2472 ACCTCATTTAGTCCATCCCGAAATTCAGTGGACCATGAC--AGGAGCCCAAGTCA 2529
 QY 2245 CTCTCGGAACCGCTCAGCTAGAGCAGACCAAGTCTTACTCAGGGCTCTCCAGCTTGT 2304
 Db 2530 CGCCTGAACAGTAGGCAFTCTCCCACTTGGGGCTAGATTGACCTTAGCCC----- 2582
 QY 2305 AGGAGCTAGGTTTACCCCTTAACCCAGGAGGACAGCTCCACCTCCAGCCCGGGAG 2364
 Db 2583 -----CAGAGCACAGTTGTACTCT-----GGAAA 2606
 QY 2365 CTTAGGACACTCAGCCCTTAGGAGTATATTTCCGCACCT----CAGAAATTCATATCT 2420
 Db 2607 CCAAGACCAATTCAGTCCCTTAGGGTAGATTTCGGCACCTTTATATAGAGTTTCACTCTT 2666
 QY 2421 CGAATCCAGCTCCCTGCCCCAATACTTCAGTCTCTCTCCAGATTTTGAATCT 2480
 Db 2667 GGAATCCAAATCCCTTACCCCAAGAACTTACATATTCTGCCCGCAATTTGGAATC-- 2724
 QY 2481 AGTTTCTCTCTCTCGTATCCCGAGTCTGGGACACAAAACCTCCGCCCGCAGCTATGAGC 2540
 Db 2725 ---TTCTCGTCTCCTAGCCCTGAGTACAGGACATGAGCTCCGACTTATCTGTGA 2781
 QY 2541 ATCTGAGCCCGCCCTCTTCTGACGAAATGCGCCCGGATCAGAGCAGGACCTCCCTT 2600
 Db 2782 ATACTGAGCTCGCCCTCTCTGACCAAACTGGCGCAAGAACAGAGCAGGACCT--CTT 2839
 QY 2601 CGACCTCTGGAACTCCAGAGGTCAGCCCATCTCGGAGCATCCCGGAGGAATCT 2660
 Db 2840 CCAGCCCTCTGGGAACCTCCCTGAGGTC-----CCAGGAGGAATCT 2881
 QY 2661 GCAGAGGGTTAGGAGTGGGTGACAAGAGCTGTATCTCTCTGTTTGTACATAGATT 2720
 Db 2882 GCACGGGTGGGGAG-----GGCCAGGCTCTCTCTGTTTGTACATAGATT 2929
 QY 2721 ATTTTTCAGTCCAGAAAGATGATACATTTTGTAAAAAATAAAAAA 2772
 Db 2930 ATTTTTCAGTCCAGAAAGATGATGATTTTGTAAAAAATAAAAAA 2981

RESULT 8
 AF016398
 LOCUS AF016398 2623 bp mRNA ROD 30-NOV-1998
 DEFINITION Mus musculus menin (Men1) mRNA, complete cds.
 ACCESSION AF016398
 VERSION AF016398.1 GI:3929517
 KEYWORDS menin.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 2623)
 REFERENCE Stewart,C., Parente,F., Piehl,F., Farnebo,F., Quincey,D.,
 Silins,G., Bergman,L., Carle,G.F., Lemmens,I., Grimmond,S.,

Xian,C.Z., Khodel,S., Teh,B.T., Lagercrantz,J., Siggers,P.,
 Calender,A., Van de Vem,V., Kas,K., Weber,G., Hayward,N.,
 Gaudray,P. and Larsson,C.
 Characterization of the mouse Men1 gene and its expression during
 development
 Oncogene 17 (19), 2485-2493 (1998)
 2 (bases 1 to 2623)
 Direct Submission
 Stewart,C.
 Submitted (29-JUL-1997) Cancer Unit, Queensland Institute of
 Medical Research, 300 Herston Road, Herston, Queensland 4053,
 Australia

FEATURES
 source 1..2623
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 gene 1..2623
 /gene="Men1"
 CDS 62..1897
 /gene="Men1"
 /codon_start=1
 /product="menin"
 /protein_id="AAC79938.1"
 /db_xref="GI:3929518"
 /translation="MGLAAOKTLPLRSIDVRLFAAELGREBPDLLVLSILGFEV
 EHLAVNRVITNPVELTPOSPADPPGLTVFPVADLSITAAIYARETAOIRGAVD
 LSLYPREGVSRRELKVKVSDVINWSLSRSFKRAHTQSLSFTGTGKLDSSGVAF
 VYVCGALGRDVLHSEDHAWVVGPNQGEAEVTHNGKNEORRGOTVWAGVER
 SWLYLKGSMRCDRMKMEAFVFCVAINPSIDLTSLQLQOQLLWLLYDLHLERY
 REALNLADLELETPGRPDPLTLYHKGISAKTYQDEHLYPYMYLAGYHCRNRV
 PMAQWADTATVIQDYNVCREDEITYKEFFEVDVIVPNLLKEASILELCEERTGE
 QAQGTQGGQSALQDPECFAHLRLFYDGLCKWEEGSPTVLHVGMATFLVQSLRGEG
 VRQKVHIVSREAAEAEPWGDREARRRGPRRESKPEEPPEPPPKKPAIDKPGSGQ
 SAGSGPPPKTSGVTGTGGQEVNAAQAPAPAPPEPPEPVLTFQSEKMGMKELL
 VATKTNSSAIKLQLTAQSQVMKKQKVSPSDYTLISFLKRRKRL"
 2608
 polyA_site /gene="Men1"
 BASE COUNT 594 a 794 c 721 g 514 t
 ORIGIN

Query Match 58.3%; Score 1616; DB 94; Length 2623;
 Best Local Similarity 79.1%; Pred. No. 5.9e-294;
 Matches 2129; Conservative 0; Mismatches 445; Indels 116; Gaps 13;

QY 90 CGCGCCGACGCGCCCGCCATGGGGCTGAAGCGCCGACGAGAGCGCTGTTCCGCTG 149
 Db 41 CGGAGGCCACCGCCCGCCATGGGGCTGAAGCGCCGACGAGAGCGCTGTTCCCTCTG 100
 QY 150 CGCTCATCGACGACGTGGTGGCCCTGTTGCTGCCGAGCTGGCGAGAGCGCGAC 209
 Db 101 CGCTCATCGACGACGTGGTGGCCCTGTTGCTGCCGAGCTGGCGAGAGCGCTGAC 160
 QY 210 CTGGTGCTCTTTCTTCTTGGTGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCTC 269
 Db 161 CTGGTGCTCTTCTTCTTGGTGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCTC 220
 QY 270 ATCCCTACCAAGTTCCTCCGAGCTCAGCTCCAGCCAGCCGCCCCCGCCGCTGGC 329
 Db 221 ATCCCCACCAAGTTCCTCCGAGCTCAGCTCCAGCCAGCCGCCCCCGCTGGT 280
 QY 330 GGCTCACCTACTTCTCCGTCGGCGACCTGTCTATCATCGCCGCTCTATGCGCGCTC 389
 Db 281 GGCTCACCTACTTCTCCGTCGGCGACCTGTCTATCATCGCCGCTCTATGCGCGCTC 340
 QY 390 ACCGCCAGATCCGAGCGCCGCTCGACCTGTCTCTCTATCTCGAGAAGGGGTGTCTCC 449
 Db 341 ACCGCTCAGATCCGCGCGCTGTGGACCTCTCTCTCTATCTCGAGAGGGAGGTGTTCT 400
 QY 450 ACCGCTGACGTGGTGAAGAGGTCTCCGATGTCTATATGGAACAGCTCAGCCGCTCTAC 509
 Db 401 AGTCGGAATCGTAAAAAAGGTCTCGGATGTCTATATGGAACAGCTCAGCCGCTCTCTAC 460

Db 1477 AGGAGATCTACAAGGAATCTTTTGAAGTGCCCAATGACGTATCCCCCAAACTGCTGAAGG 1536
Qy 1243 AGGCAGCAGCTCTGCTAGGCGGGCAGGAGCGCGGGGAGCAAGACGAGGCACCC 1302
Db 1537 AGGCGGCAGCTGCTAGAAACAGGCGAAGAGCGGACTGGGGAGCAAGCCAGGGCACCC 1596
Qy 1303 AGAGCCAAAGGTTCCGCCCTCCAGGACCCCTGAGTGTCTTTCGCCCACTGCTCGGATTTACG 1362
Db 1597 AGGSCCAAGGTTCTGCTCTCCAGGACCCAGAAATGCTTTTGGCCCACTGCTCGGATTTATG 1656
Qy 1363 AGGCATCTGCAATGAGGAGGAGGAGTCCAGGCGCTGTGCTGACGTGGGCTGGGCCA 1422
Db 1657 ATGCACTCTGCAATGAGGAGGAGGTAGCCCCACACCACTGACTGCAGCTAGGCTGGGCAA 1716
Qy 1423 CTTTCTTGTGTCAGTCTCCTAGGCGGTTTTGAGGGACAGGTGCGGCAGAAAGTGCATAG 1482
Db 1717 CTTTCTTGTGTCAGTCTCCTAGGCGGCTTCGAGGGACAGGTGCGGCAGAAAGTGCATAG 1776
Qy 1483 TGAGCCGAGAGCGGAGCGGCCGAGGCGCTGGGGCGAGGAAGCCCGGGAAG 1542
Db 1777 TTACCCGGGAGACAGAGCGGCCGAGCTGAAGAACCATGGGGGATGAAGCCGAGAAG 1836
Qy 1543 GCCGGCGGGGGCCACAGCGGGGAGTCCAAAGCCAGAGGAGCCCGCCGCCCAAGAAC 1602
Db 1837 GCCGTGCGGCTGTGCTCCCGAAGAGAGTCCAAAGCCGTGAGGAGCCACCAACCCAGAAAGC 1896
Qy 1603 CAGCACTGGACAAGGGCTTGGGCAACCGGCCAGGCTGAGTGTCAAGGACCCCGCGGAAGC 1662
Db 1897 CTGCAATTTGGACAAAGGGCCGGCTCAGGACAAAGTGCAGGGTCCGGGACCACTAGGAAAA 1956
Qy 1663 CTCCTGGGACTGTGCTGGCACAGCC - - - GAGGCCCTTGAAGTGGGACGACCGCTCAGG 1719
Db 1957 CGTCAGGAGTGTCCCAGGTACTACCCGCGGTGCGCCAGGAAGTCGGCAATGCTGCTCAGG 2016
Qy 1720 TGCAGACACCGCAGCATCACCCCGGGAGGCTCAGTGTCTACCTTTCCAGAGTGAGA 1779
Db 2017 CTCAGACACTTGCAGCATCGCCACCCCGAGGCGCCAGTGTCTACCTTTCCAGAGTGAGA 2076
Qy 1780 AGATGAAGGATCAAGGAGTGTGCTGGCCACCAAGATCAACTCGAGGCGCCATCAAGC 1839
Db 2077 AGATGAAGGATCAAGGAGTGTGCTGGCCACCAAGATCAACTCGAGGCGCCATCAAGC 2136
Qy 1840 TGCAACTCAGGCGCAGTTCGCAAGTGCAGATGAAGAGACGAAAGTGTCCACCCCTAGTG 1899
Db 2137 TGCAACTCAGGCGCAGTTCGCAAGTGCAGATGAAGAGACGAAAGTGTCCACCCAGCGC 2196
Qy 1900 ACTACACTCTGCTTCTTCTCAAGCGGCGAGCGCAAGGCTCTGAACCTACTGGGCACTTCG 1959
Db 2197 ACTACACTCTTCTTCTTCAAGCGGCGAGCGCAAGGCTCTGA - - - - - ACTCTG 2246
Qy 1960 GACGCGTTGTGGGACCGAGCTCCGCTTGTAGTCCCGCAACTCTGAGCGCCATGTTCTGCC 2019
Db 2247 ATTCTGTGCAAGGCCCGCAGGCTGCGCTCACTCCAGAGCGCTTCACTCCCTCTA 2306
Qy 2020 CCCAGCCCAAGGGGACAGGCGCTCACTCTACCAAAACCTAGTGTCCGCTCCCGAGTA 2079
Db 2307 ATTCAAGCCTAGAC - TCGCGCTCAGTCTAGGCGGACCCAGAGCTGCTCAGTGACTA 2365
Qy 2080 CAGTCTGTATCAAAACCCAGATTTTCTCCAGCTCAGAACCCAGGCTGTGCCCGACGCT 2139
Db 2366 G - - - - - GAGCCTTCAATTTAGCCCTCTTCAGGCGCCCGAGAGCCCTGCTCAGTTCT 2415
Qy 2140 TAGAATATAGTCTCTTCTCCAGAAATCCCGCGGCGCAATGGAACCTCTACGCTGGGTC 2199
Db 2416 TAGAATATAGTCTCTTCTCCAGAAATCCCGCGGCGCAATGGAACCTCTATTTAGTGTG 2459
Qy 2200 CTAATTACCACTCTTTAAAGGCCCGAGCCCTTAGAACCCAGCTCTCTCCGGAACCGCT 2259
Db 2470 CAATCCCCCGAAATTCAGTGGACCATGAC - - - AGGAGCCCAAGTACCGCTCAGACAGTAGG 2527
Qy 2260 CACCTAGAGCCGACCAACGCTTACTCAGGCGCTCTCCAGCTTGTAGGAGCTGAGGTTTC 2319
Db 2528 GCATTCCTCCCACTTGGGGGCTAGATTGACCTTAGGCC - - - - - 2565

Qy 2320 ACCCTTAACCAAGGAGGACACAGTCCCACCTCCAGCCCGGGAGGCTTAGGACCACTTCAG 2379
Db 2566 - - - - - CAGAGCACAGTTGCTACT - - - - - GGAACCAAGGACCACTTCAG 2604
Qy 2380 CCCTAGGAGTATATTTCCGCACTT - - - - - CAGAAATTCATATCTTGGCAGTCCAGCTCCCT 2437
Db 2605 TCCTAGGGGTAGATTTCCGCACTTATAGAGTTCCATCTTTTGGGAATCCAAATCCCTT 2664
Qy 2438 GCCCAAAATCACTCAGTCTCTCCAGAAATTTGGAAATCTTAGTTCCTCTCTCTTCGT 2497
Db 2665 ACCCAAAAGAACCTTACATATTTGCTGCGGAAATTTGGAATC - - - - - TTCCTCTCTCTTCCT 2719
Qy 2498 ATCCGAGTCTGGGACACAAACTCCGCCCGCCAGCTTATGAGCATCTGAGCCCGCCCT 2557
Db 2720 AGCCCTGAGTACAGGACATGAAGCTCCGACTCTAATCTCTGAATAGCTCGGCC 2779
Qy 2558 CTTCCTGACAAACTGGCCCGGATCAGAGCAGGACCTCTCCGACCTTCCGACCTTTCGGGAAC 2617
Db 2780 CTTCCTGACAAACTGGCGCAAGAACAGAGCAGGACCT - - - - - CTTCAGCCTCTTGGGAAC 2837
Qy 2618 TCCAGAGTCCAGCCCATCTCGGAGCATCCCGGAGGAAATCTGCAGAGGGTTAGGAGT 2677
Db 2838 TCCTGAGGTC - - - - - CCAGAGGAAATCTGCACCGGTTGGGGAG - 2878
Qy 2678 GGTGACAGAGGCTGATCTCTCTCTTTTGTACATAGATTTATTTTTCAGTTCCTCAAGA 2737
Db 2879 - - - - - GGCCAGCCTCTCTCTTTGTACATAGATTTATTTTTCAGTTCCTCAAGA 2927
Qy 2738 AGATGAATACATTTTGTGA 2757
Db 2928 AGATGAATACATTTTGTGA 2947

RESULT 10

AF072755 2577 bp mRNA ROD 11-JUL-1998

LOCUS Mus musculus MENIN (Men1) mRNA, complete cds.

DEFINITION AF072755

ACCESSION AF072755

VERSION AF072755.1 GI:3309058

KEYWORDS house mouse.

SOURCE Mus musculus

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

REFERENCE 1 (bases 1 to 2577)

AUTHORS Bassett,J.H.D. and Thakker,R.V.

TITLE cDNA sequence of mouse MEN1 gene

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 2577)

AUTHORS Bassett,J.H.D. and Thakker,R.V.

TITLE Direct Submission

JOURNAL Submitted (15-JUN-1998) MRC Molecular Endocrinology, Imperial

College School of Medicine, Hammersmith Hospital DuCane Road,

London, NA W12 0NN, United Kingdom

FEATURES

Location/Qualifiers

1..2577

/organism="Mus musculus"

/db_xref="taxon:10090"

/chromosome="19"

/notes="map location based on human syntony"

1..2577

/gene="Men1"

46..1881

/gene="Men1"

/notes="similar to Homo sapiens MENIN"

/codon_start=1

/product="MENIN"

/protein_id="AAC26001.1"

/db_xref="GI:3309058"

/translation="MGLKAAQKTLFPLRSIDDDVVRFLPAELGREEDPIVILSILVIGFV

EHLAVNRVITNPVLPSTFPPADLSITIAALYARFPAQIRGAVD

LSLYPREGGVSSRELKVKVSDVWNSLSRSFYFKDRAHIQSLFSFITCTKLDSSGVAF

VVGAQALGLRDVHLAISEDHAWVFGPNGEQTAEVTHWKGKNEEDRRQTVNAGVAER
SWLYLKGSMYRCDKMEVAFMVCAINPSIDLHTDSLEYDLOLQQLWLLYLGLHLEARY
PMALGNLADLELEPTGPPDPPLTYLHKGIASAKTYOYDEHYIPVYLAGYHCRNRNV
REALQAWADTATVODNYNCRDEDELYKEFEVANDVTPNLLKKAASILETGERBTE
QAQGTQGGSLADQPECFYHALLRYPYDGLCKWEEGSPVPLVHGWNATFLVQSLRFEGQ
YRQKVMVSRARAGAEPEWMDARERRRGRPRRESKPEPPPPKPKPALDKGPGSGQ
SAGLPPRPSTGTVPTGTTRGGOEVGNAAQAPAPASPPPEPVLTFQSEKMKMKMKELL
VATIRNSAIKLOLTAQSOQVMKOKVSTPSDYTLSELKQRKGL"
BASE COUNT 574 a 785 c 712 g 506 t
ORIGIN

Query Match 57.2%; Score 1585; DB 94; Length 2577;
Best Local Similarity 78.7%; Pred. No. 4e-288;
Matches 2090; Conservative 0; Mismatches 450; Indels 115; Gaps 12;

QY 90 CGCGCCGACCGCCGCGCCATGGGCTGAAGCGCGCCAGAACGCGTGTCCGCTG 149
DB 25 CGAGGCCACCGCCCGCCCGCTGAAGCGCGCCAGAACGCGTGTCCCTG 84
QY 150 CGCTCCATCGAGAGTGGTGGCCCTGTTTGTGCTGCGAGCTGGGCGGAGGAGCGGAC 209
DB 85 CGCTATCGAGAGTGGTGGCCCTGTTTGTGCTGAGAGCTGGGCGGAGGAGCGCTGAC 144
QY 210 CTGGTGTCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 269
DB 145 CTGGTGTCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTCCTTTC 204
QY 270 ATCCCTACCAAGTTCCTGAGCTACCTTCCAGCCAGCCGCGCCGCGCCGCGCTGCG 329
DB 205 ATCCCAACCAAGTTCCTGAGCTACCTTCCAGCCAGCCGCGCCGCGCCGCGCTGCGT 264
QY 330 GGCCTCACCTACTTTCCTGCGGACCTGTCTATCATCGCGCCCTCTATGCGCGCTTC 389
DB 265 GGCCTCACCTACTTTCCTGCGGACCTGTCTATCATCGCGCCCTCTATGCGCGCTTC 324
QY 390 ACCGCCAGATCCGAGGCGCTCGACCTGTCCTCTATCTCGAAGAGGGGTGTCCTCC 449
DB 325 ACCGCTCAGATCCGCGGCGCTGTGGACCTCTCCCTCTATCTCGAGAGGGGTGTTCT 384
QY 450 AGCGGTGAGCTGTGAAGAGCTCTCCGATGTCATATGAACAGCTCAGCGCTCCTAC 509
DB 385 AGTCGGAACTGGTAAAGAGCTCTCGGATGTCATATGAACAGCTCAGCGCTCCTAC 444
QY 510 TTCAAGGATCGGGCCACATCCAGTCCCTCTTTCAGTTCATCACAGGACCAAAATGGAC 569
DB 445 TTCAAGGACCGGCGCCACATCCAGTCCCTCTTTCAGTTCATCACAGGACCAAAATGGAC 504
QY 570 AGCTCGGTGTGGCTTTGCTGTGGTGGGCTGCGGAGGCGCTGCGGAGGATGTC 629
DB 505 AGCTCGGGGCTGGCTTTGCTGTGGTGGGCTGCGGAGGCTGCGGAGGATGTC 564
QY 630 CACCTCGGCTCTGAGGATCATGCTGGTGTGGTGTGGGCGCCAAATGGGAGCAGACA 689
DB 565 CATCTGGCCCTCTGAGGATCATGCTGGTGTGGTGTGGGCGCCAAATGGGAGCAGACA 624
QY 690 GGTGAGGTCACTTGGCAGCGGCAAGGACGAGGACCGGAGGCGGAGGACAGTCAATGCG 749
DB 625 GCTGAGGTGACGTGGCAGCGGCAAGGACGAGGACCGGAGGCGGAGGACAGTCAATGCG 684
QY 750 GGTGTGGCTGAGCGAGTGGTGTACCTGAAAGGATCATATCATGCGTGTGACCGCAAG 809
DB 685 GGTGTGGCTGAGCGAGTGGTGTACCTGAAAGGATCATATCATGCGTGTGACCGCAAG 744
QY 810 ATGGAGGTGGCTTATGTGTGTGGCTATCAACCTTTCATGACCTGCACACCGACTCG 869
DB 745 ATGGAGGTGGCTTATGTGTGTGGCTATCAACCTTTCATGACCTTTCATGACCTGACTCT 804
QY 870 CTGGAGCTTCTGAGCTGCAGCAGAGTGTCTGGCTCTCTATGACCTGGGAGATCTG 929
DB 805 TTGGAGCTTGTGCACTACAGCAGAGTGTCTGGCTCTCTATGACCTCGGACATCTG 864
QY 930 GAAAGGTACCCCATGGCTTTAGGGAACCTTGGCAGATCTTAGAGAGGCTGGAGCCCAACCCCT 989

DB 865 GAAAGATACCCCATGGCGCTAGGGAACCTTGGCAGACCTGGAGAGCTGAGCCTACCCCG 924
QY 990 GCGCGCCAGACCCCACTACCCCTCTACCAACAAGGCAATTTGCCTCAGCAAGACCTACTAT 1049
DB 925 GCGCGCCAGACCCCACTACCCCTTTATCAAGGGAATTTGCCTCAGCTAAGACCTACTAC 984
QY 1050 CGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTGTGCGAACCCCAAT 1109
DB 985 CAGGATGAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTGTGCGAACCCCAAT 1044
QY 1110 GTGCGGGAAGCCCTGAGGCTTGGGCGGACACGGCCACTGTTCATCCAGACTACAATAC 1169
DB 1045 GTGCGGAAGCCCTGAGGCTTGGGCGGACACTGCCACTGTTCATCCAGACTACAATAC 1104
QY 1170 TCGCGGAGACGAGAGATCTACAGGAGTCTTCTTGAAGTAGCCCAATGATGTCATCCCC 1229
DB 1105 TCGCGGAGGATGAGGAGATCTACAGGAAATCTTTTGAAGTGGCAATGACGTTCATCCCC 1164
QY 1230 AACCTGCTGAAGGAGGACGACGCTTGTCTGAGCGGCGGAGGAGCGCGGGGAGCA 1289
DB 1165 AACCTGCTGAAGGAGGCGGCGGACGCTGTAGAAACAGGCGAAGAGCGGACTGGGAGCA 1224
QY 1290 AGCCAGGACCCAGAGCAAGGTTCCGCCCTTCCAGGACCTTGAGTGTCTCGCCCACTG 1349
DB 1225 GCGCAGGACCCAGGCGCAAGGTTCTGCTCTCCAGGACCCAGAAATGCTTTGCCCACTG 1284
QY 1350 CTGCGATTCTACGAGGCACTGCAAAATGAGGAGGAGGAGTCCCAAGCTGTGCTGCAC 1409
DB 1285 CTGCGATTCTTATGATGGCATCTGCAAAATGGGAGGAGGTTAGCCCAACAGCTGTGCAC 1344
QY 1410 GTGGGTGGGCGACCTTCTTGTGAGTCCCTTAGGCGCTTTTGAAGGACAGGTGCGGCGAG 1469
DB 1345 GTAGGTGGGCAACTTTCCTTGTGAGTCCCTTAGGCGCTTTCAGGAGACAGGTGCGGCGAG 1404
QY 1470 AAGTGTGCGCATGTGAGCGAGAGCGGCGGCGGAGGCGGCGGAGGCGGTGGGGCGAG 1529
DB 1405 AAGGTGCATGTGTAGCGGGAAGCAGAGGCGGCGGCGGCTGAAGAACCATGCGGGGAT 1464
QY 1530 GAAGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGAGTCCCAAGCCAGAGAGCCCGCG 1589
DB 1465 GAAGCGCGGAGGCGGCGGCGGCGGCGGCGGCGGAGTCCCAAGCCAGAGAGCCCGCA 1524
QY 1590 CGCGCCCAAGAACCCAGCACTGGACAGGCGCTTGGGCGGCGGCGGCGGAGTGTGTCAGA 1649
DB 1525 CCACCCCAAGAACCTTGCATTGGACAGGCGCGGCGGCTCAGGACAAAGTGCAGGTTGGGA 1584
QY 1650 CCGCCCGGGAAGCTGCTGGGACTGTGCTGGGCGGCGGCGGCGGCGGAGTGAAGTGGC 1706
DB 1585 CCACCTAGGAAACGCTCAGGGACTGTGCTGGGCGGCGGCGGCGGAGTGAAGTGGC 1644
QY 1707 AGCAGGCTCAGGTGCGGCGGCGGCGGCGGCGGCGGAGTCCCAAGCCAGAGGCTCCAGTGCAC 1766
DB 1645 AATGCTGCTCAGGCTCCAGCACTTGCAGGATCGCCACCGGCGGCGGCGGAGTGCAC 1704
QY 1767 TTCCAGAGTGAAGAAGTGAAGGCGATGAAGGAGCTGCTGGTGGCGGCGGCGGAGTCAACTCG 1826
DB 1705 TTCCAGAGTGAAGAAGTGAAGGCGATGAAGGAGTACTGGTGGCGGCGGCGGAGTCAACTCG 1764
QY 1827 AGCGCATCAAGCTGCAACTCAGGCGGCGGCGGCGGCGGAGTGCAGATGAAGAGCAAGAGTG 1886
DB 1765 AGCGCATCAAGCTGCAACTCAGGCGGCGGCGGCGGCGGAGTGCAGATGAAGAGCAAGAGTG 1824
QY 1887 TCCACCCCTAGTGAAGTGAAGTGAAGGCGGCGGCGGCGGAGTGCAGATGAAGAGCAAGAGTG 1946
DB 1825 TCCACCCAGGAGTGAAGTGAAGTGAAGGCGGCGGCGGCGGAGTGCAGATGAAGAGCAAGAGTG 1881
QY 1947 ACTGGGAGTTCGGGAGTTCGGGAGTTCGGGAGTTCGGGAGTTCGGGAGTTCGGGAGTTC 2006
DB 1882 -----ACTCTGATCTGTGGCAAGGCGGCGGCGGCGGAGTTCGGGAGTTCGGGAGTTC 1934
QY 2007 CCCATGTTCTGCCCGGCGGCGGCGGCGGCGGAGTTCAGCTTCTTACCCCAAGCCCTAGGTTTC 2066

Qy	87	GGCGCGCGCCACGCGCGCGCCATG	GGGCGTGAAGCGCGCCAGAA	GACGCTGTTCCGG	141
Db	26	GCCCGAGCGACGCGCGCGCCATG	GGGCGTGAAGCGCGCCAGAA	GACGCTGTTCCGT	85
Qy	147	CTGGCGTCCATCAGCAGCGTGGT	GGCGCTGTTGCTGGCGGAGAGCG	GCTGGCGGAGAGCGG	206
Db	86	CTGGCGTCTATCAGCAGCGTGGT	GGCGCTGTTGCTGGCAGAGCTGG	CGCGAGAGAGGCT	145
Qy	207	GACCTGCTGCTCTTTCCTTGGT	GGCGTCTGTTGGACATTTTCT	TGGCTGTCAACGCG	266
Db	146	GACCTGCTGCTCTTTCCTTGGT	GGCGTCTGTTGGACATTTTCT	TGGCTGTCAACGCT	205
Qy	267	GTCAATCCCTACCAACGTTCCG	AGCTACCTCCAGCCCGCGCGCG	CGCGCGCGCGCGCT	326
Db	206	GTCAATCCCAACGTTCCGAGCT	ACCTTCCAGCCCGCGCGCGCG	ACCGAGCGCTCT	265
Qy	327	GGCGGCTCACTTCTTCCTGG	CGCGACCTGTCTATCATCG	CGCGCTCTATCGCGCG	386
Db	266	GGTGGCTCACTTCTTCCTGG	CGCGACCTTATCATCGCGCG	CTCTATCGCGCG	325
Qy	387	TTCACCGCCCGAGATCCGAG	CGCGCTCGACCTGTCCCTCT	ATCCTCGAAGGGGGTCTC	446
Db	326	TTCACCGCTCAGATCCGCG	CGCTCTCGACCTCTCTCTCT	ATCCTCGAGGGAGGTGT	385
Qy	447	TCCAGCGCTGAGCTGGTGA	AGAAGTCTCCGATGCTATAT	GAAACAGCCTCAGCGCTCC	506
Db	386	TCTAGTCGCAACTGGTAA	AAAAAGCTTCGGATGCTAT	ATGAAACAGCCTCAGCGCTCC	445
Qy	507	TACTTTAAGGATCGGGCC	ACATCCAGTCCCTCTTCAG	CTTCATCAGAGCACCAATTG	566
Db	446	TACTTTAAGGACCGGGCC	ACATCCAGTCCCTCTTCAG	CTTCATCAGAGCACCAACTG	505
Qy	567	GACAGCTCCGGTGTGGCT	TTTGTGTGGTGGCGCTGCC	AGCGCTGGGTCTCGGGAT	626
Db	506	GACAGCTCGGGCGTGGCT	TTTGCAGTGGTAGGGCGCT	GCGACGCTCTGGGTCTCAGAT	565
Qy	627	GTCCACCTCGCCCTGTCT	GAGGATCATGCTGGGTAGT	TTTGGGCCCCAATGGGAGCAG	686
Db	566	GTCCATCTGGCCCTGTCT	GAGATCATGCTGGGTGGT	TTTGGGCCCCAACGGGAGCAG	625
Qy	687	ACAGCTGAGGTCACTGGC	ACGCAAGGGCAACGAGGAC	CGCAGGGGCCAGACAGTCAAT	746
Db	626	ACAGCTGAGGTGAGTGGC	ACGCAAGGCAAGGACCG	AGGACCGCAGAGCCAGACAGTCAAT	685
Qy	747	GCCGGTGTGGCTGAGCG	GAGCTGGCTACCTGAAAG	AGATCATATGCGCTGTGACCGC	806
Db	686	GCCGGTGTGGCTGAGCG	GAGCTGGCTACCTGAAG	GGTCTGATGCGCTGGACCGT	745
Qy	807	ANGATGAGGTGGCGTTCAT	GTGTGCCATCAACCTTCCAT	TGACCTGCACACCGAC	866
Db	746	ANGATGAGGTGGCGTTCAT	GTGTGCCATCAACCTTCCAT	GTGATCTTCACACTGAC	805
Qy	867	TCGCTGGAGCTTCTGCA	GCTGCAGCAGAGCTGCT	CTGGCTGTCTATGACCTGGAGAT	926
Db	806	TCCTTTGAACTGTTGCA	CTTACAGCAGAGCTGCT	CTGGCTGTATGACCTCGGACAT	865
Qy	927	CTGAAAGGTACCCCATGG	CGCTTTAGGCAACCTGGC	AGATCTAGAGAGCTGGAGCCACC	986
Db	866	CTGAAAGATACCCATGG	CGCTTAGGCAACCTGGC	AGATCTAGAGAGCTGGAGCGCTACC	925
Qy	987	CTGGCGCGCGACGAC	CTACCTCTTACCACAAG	GGGATTTGCCCTCAGCCCAAGACCTTAC	1046
Db	926	CCGGCGCGCGACGAC	CTACCTCTTATCACAAG	GGGATTTGCCCTCAGCTAAGACCTTAC	985
Qy	1047	TATCGGATGAACAGAT	CTACCTCTACCTGCTGG	CTTACACTGTGCGCAGCGC	1106

Query Match	54.8%;	Score	1518.2;	DB	94;	Length	1906;
Best Local Similarity	88.3%;	Pred. No.	1.5e-215;				
Matches	1649;	Conservative	0;	Mismatches	218;	Indels	0;
Gaps							
0;							
QY	90	CGCGCGCCACCGCGCGCGCCATGGGGCTCAAGGCCCGCCAGAAAGACGCTGTTCGCCGCTG	149				
DB							
	2	CGAGGCCACACGCGCGCGCCATGGGGCTGAAGGCCCGCCAGAAAGAGCGCTGTTCCTCTGT	61				
QY	150	CGCTCATCGACGACGTGTGCGCTGTTTTCGTCGCCAGCTGGCCCGAGAGACCCGGAC	209				
DB							
	62	CGCTCATCGACGACGTGTGCGCTGTTTCGCGAGCTGGGCCGAGAGAGCCGTGAC	121				
QY	210	CTGGTGTCTCTTTCCTGGTGCTGGGCTTCGTGGAGCATTTTCGTGGCTGTCAACCGGGTC	269				
DB							
	122	CTGGTGTCTCTCTCTGGTTCGTGGGCTTCGTAGAGCATTTCTTGGCTGTCAACCGGTCT	181				
QY	270	ATCCCTACCAACGTTCCCGAGCTCACCTTCAGCCCAAGCCCGCCCGCCGACCCGCTGGC	329				
DB							
	182	ATCCCAACCAACGTGCGGAGCTCACCTTCAGCCCAAGCCCGCCGACCCGCTTCCTGGT	241				
QY	330	GGCCTCACCTACTTTCGCTGGCGGACCTGCTATCATCGCCGCCCTCTATGCCCCGCTTC	389				
DB							
	242	GGCCTCACCTACTTCCAGTGGCGGACCTATCCATCATTTGCTGCCCCCTATGCCCCGATT	301				
QY	390	ACGCCCAAGATCCGAGGCGCGCTGCGACCTGTCCCTCTATCTCTCGAAGAGGGGTGTCTCC	449				


```
Db 8399 ACTCAGGGCTCCTCCAGCTGTAGGAGCTGAGGTTTACCCCTTAACCCAGGAGCACA 8458
QY 2342 GGTCCACCTCCAGCCGGGAGGCTAGGACCACTAGGCCCTAGGAGTATATTTCCGCA 2401
Db 8459 GGTCCACCTCCAGCCGGGAGGCTAGGACCACTAGGCCCTAGGAGTATATTTCCGCA 8518
QY 2402 CTTTCAGAAATTCATATCTTCGGAATCCAAGCTCCCTGCCCAAACTACTTCAGTCTGCT 2461
Db 8519 CTTTCAGAAATTCATATCTTCGGAATCCAAGCTCCCTGCCCAAACTACTTCAGTCTGCT 8578
QY 2462 TCCAGAAATTTGGAATCCTAGTTCTCTCTCTCGTATCCCGAGTCTGGGACACAAACT 2521
Db 8579 TCCAGAAATTTGGAATCCTAGTTCTCTCTCTCGTATCCCGAGTCTGGGACACAAACT 8638
QY 2522 CCGCCCGCCAGCTATGAGCACTCTGAGCCCGCCCTCTCTCTGACGAAACTGCGCCCGGA 2581
Db 8639 CCGCCCGCCAGCTATGAGCACTCTGAGCCCGCCCTCTCTCTGACGAAACTGCGCCCGGA 8698
QY 2582 TCAGAGCAGGACCTCCCTCCGACCCCTCTCGGAACTCCAGAGGTCAGCCCATCTCGG 2641
Db 8699 TCAGAGCAGGACCTCCCTCCGACCCCTCTCGGAACTCCAGAGGTCAGCCCATCTCGG 8758
QY 2642 AGATCCCGAGGAAATCTCCAGAGGGTTAGGAGTGGTGACAAGAGCTGATCTCTTC 2701
Db 8759 AGATCCCGAGGAAATCTCCAGAGGGTTAGGAGTGGTGACAAGAGCTGATCTCTTC 8818
QY 2702 CTGTTTGTACATAGATTATTTTTCAGTTCCCAAGAAAGATGAATCATTTTGTAAAAA 2761
Db 8819 CTGTTTGTACATAGATTATTTTTCAGTTCCCAAGAAAGATGAATCATTTTGTAAAAA 8878
QY 2762 AAAAAAAA 2770
Db 8879 AAATATAA 8887

RESULT 15
AC000134/c 203300 bp DNA PRI 06-MAY-1999
LOCUS Homo sapiens Chromosome 11q13 BAC Clone 137c7, complete sequence.
DEFINITION AC000134
ACCESSION AC000134
VERSION AC000134.14 GI:4755212
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 203300)
AUTHORS Crabtree, J. and Roe, B.A.
TITLE sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 2 (bases 1 to 203300)
AUTHORS Chandrasekharappa, S.C., Guru, S.C., Manickam, P., Olufemi, S.-E. and
Collins, F.S.
TITLE Sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 3 (bases 1 to 203300)
AUTHORS Emmert-Buck, M.R., Debelenko, L.V., Zhuang, Z., Lubensky, I.A. and
Liotta, L.A.
TITLE Sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 4 (bases 1 to 203300)
AUTHORS Weisemann, J. and Boguski, M.S.
TITLE Sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 5 (bases 1 to 203300)
AUTHORS Agarwal, S.K., Kester, M., Kim, Y.S., Heppner, C., Dong, Q.,
Spiegel, A.M., Burns, L.A. and Marx, S.J.
TITLE Sequence of a 11q13 bac mapping to PYGM
JOURNAL Unpublished (1997)
REFERENCE 6 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
```

```
JOURNAL Submitted (12-MAR-1997) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 7 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (23-MAR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 8 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (24-APR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 9 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (30-APR-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 10 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 11 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (04-MAY-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 12 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (05-MAY-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
REFERENCE 13 (bases 1 to 203300)
AUTHORS Roe, B.A.
TITLE Direct Submission
JOURNAL Submitted (06-MAY-1999) Department Of Chemistry And Biochemistry,
The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman,
OK 73019, USA
COMMENT On May 6, 1999 this sequence version replaced gi:4753208.
FEATURES
source 1..203300
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="137c7"
BASE COUNT 49152 a 57085 c 53982 g 43081 t
ORIGIN
Query Match 45.2%; Score 1253.2; DB 85; Length 203300;
Best Local Similarity 96.2%; Pred. No. 3e-226;
Matches 1337; Conservative 0; Mismatches 48; Indels 5; Gaps 5;
QY 1382 GGAGGGCAGTCCACGCCCTGCTGCTGCACGTGGGCTGGGCCACCTTTCTTGTGAGTCCCT 1441
Db 156588 GGATGGCCAGAGCAGGGTCTTGAGTTCCAGCCACTGGCGCGCAACCTTGCTCTCACCTT 156529
QY 1442 AGGCCGTTTTGAGGGACAGAGTGGCGGAGAGGTGCCCATAGTCGAGCGAGAGCCGAGGC 1501
Db 156528 GCTCTCCCACTGGCCAGGTGGCGGAGAGGTGCCCATAGTCGAGCGAGAGCCGAGGC 156469
QY 1502 GGCCGAGGCGGAGGAGCCGTGGGGCGGAGGAGCCCGGAGAGCGGCGGCGGCCACG 1561
Db 156468 GGCCGAGGCGGAGGAGCGGTGGGGCGGAGGAGCCCGGAGAGCGGCGGCGGCCACG 156409
QY 1562 GCGGAGTCCAAAGCCAGAGGAGCCCGCGGCCCAAGAGCCACTGCACAGGCGCT 1621
```

Db 156408 GCGGGAGTCCAAGCCAGAGAGAGCCCCCGCCGCGCCCAAGAACCCAGCAGCTGGACAAGGGCCT 156349
QY 1622 GGGCACCAGGCTGAGTGTACAGGACCCCGCCGGAAGCCTCTCGGAGCTGTGCTGG 1681
Db 156348 GGGCACCAGGCTGAGTGTACAGGACCCCGCCGGAAGCCTCTCGGAGCTGTGCTGG 156289
QY 1682 CACAGCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCCGCGAGCATCAC 1741
Db 156288 CACAGCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCCGCGAGCATCAC 156229
QY 1742 ACCGGCGAGGCTCAGTGTCTCACTTTCCAGAGTGAGAAGATGAAGGGCATGAAGAGCT 1801
Db 156228 ACCGGCGAGGCTCAGTGTCTCACTTTCCAGAGTGAGAAGATGAAGGGCATGAAGAGCT 156169
QY 1802 GCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCA 1861
Db 156168 GCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCA 156109
QY 1862 AGTGAGATGAAGAGCAGAAAGTGTCCACCCTAGTGACTACACTCTGTCTTTCTCTCAA 1921
Db 156108 AGTGAGATGAAGAGCAGAAAGTGTCCACCCTAGTGACTACACTCTGTCTTTCTCTCAA 156049
QY 1922 GCGGAGCGCAAGGCCCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCAGGC 1981
Db 156048 GCGGAGCGCAAGGCCCTGAACCTACTGGGACTTCGGACCGCTTGTGGGACCCAGGC 155989
QY 1982 TCCG-CCTTAGTCCCGCCAACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGC 2040
Db 155988 TCCGCGCTTAGTCCCGCCAACTCTGAGCCCATGTTCTGCCCGCCAGCCCAAGGGGACAGGC 155929
QY 2041 CTCACCTCTACCCAAACCCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGA 2100
Db 155928 CTCACCTCTACCCAAACCCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGA 155869
QY 2101 TTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTTAGAAATATAGGTCTCTTCC 2160
Db 155868 TTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTTAGAAATATAGGTCTCTTCC 155809
QY 2161 CAGAATCCAGCGCGCCCAATGGAACCTCACGCTGGGTCTTAATACCACTCTTTAAAGG 2220
Db 155808 CAGAATCCAGCGCGCCCAATGGAACCTCACGCTGGGTCTTAATACCACTCTTTAAAGG 155749
QY 2221 CCCAGCCCTAGAAACCCCAAGCTCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGT 2280
Db 155748 CCCAGCCCTAGAAACCCCAAGCTCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGT 155689
QY 2281 TACTCAGGCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCCAGGAGCAC 2340
Db 155688 TACTCAGGCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCCCAA-GGAGCAC 155630
QY 2341 AGGTCCCACTCCAGCCCGGGAGGCTTAGGACCACTCAGCCCTAGGAGTATATTTCCGC 2400
Db 155629 AGGTCCCACTCCAGCCC-GGAGGCTTAGGACCACTCAGCCCTAGGAGTATATTTCCGC 155571
QY 2401 ACTTCAGAATTCATATCTTGGAAATCCAAGCTCCCTGCCCAATAAAGTTCAGTCTCTGC 2460
Db 155570 ACTTCAGAATTCATATCTTGGAAATCCAAGCTCCCTGCCCAATAAAGTTCAGTCTCTGC 155511
QY 2461 TTCCAGAATTTGGAAATCCTAGTTTCTCTCTCTCGTATCCCGAGTCTGGGACACAAAC 2520
Db 155510 -TCCAGAATTTGGAAATCCTAGTTTCTCTCTCTCGTATCCCGAGTCTGGGACACAAAC 155452
QY 2521 TCCGCCCCAGCCTATGAGCATCTGAGCCCGCCGCTCTTCTGAGCAACTGGCCCGG 2580
Db 155451 TCCGCCCCAGCCTATGAGCATCTGAGCCCGCCGCTCTTCTGAGCAACTGGCCCGG 155392
QY 2581 ATCAGAGCAGGACCTCCCTTCCGACCTCTTGGAACTCTCCAGAGTCCAGGCCATCTCG 2640
Db 155391 ATCAGAGCAGGACCTCCCTTCCGACCTCTTGGAACTCTCCAGAGTCCAGGCCATCTCG 155332
QY 2641 GAGCATCCGGAGGAATCTCGAGAGGGTTAGGAGTGGGTGACAGAGCCTGATCTCTT 2700
Db 155331 GAGCATCCGGAGGAATCTCGAGAGGGTTAGGAGTGGGTGACAGAGCCTGATCTCTT 155273

QY 2701 CCTGTTTCTGATAGATTTATTTTCACTTCCCAAGAAAGATCAATACATTTTG1 16AAA 2760
Db 155272 CCTGTTTCTGATAGATTTATTTTCACTTCCCAAGAAAGATCAATACATTTTGT1AAAA 155213
QY 2761 AAAAAAAA 2770
Db 155212 AAAAAATAAA 155203

Search completed: October 14, 2001, 01:32:59
Job time: 20395 sec

THIS PAGE BLANK (USPTO)


```
Db 1 MGIHQKHHFPLRGIDGVQVLFSELSNPEPDLALLSLVLGFEHFLAVNRVVPVNPVG 60
QY 61 LTFQSPADPPGGLTYFPVADISITAAIYARFTAQIRGAVDLSLYPREGSVSSRELVK 120
Db 61 VFEPL--OPSLNSCFTFVLLNLSALYERFMAQIRGAVDMQSRKPSGASSRELVK 117
QY 121 VSDVWNSLSRSYFKDRAHIQSLFSPITCKLDSSGAVAFVVGACQALGLRDVHLALSED 180
Db 118 VSDVWNSLSRSYFKDRAHIQSLFSLITCTKLDSSGAVAFVVGACQALGLRDVHLALSED 177
QY 181 HAWVPGNGEQTAEVTHWKGNEDERRGQTVNAGVAERSWLYLKGSMCRDKMEVAFV 240
Db 178 HAWVPGNGEQTAEVTHWKGNEDERRGQTVNAGVAERSWLYLKGSMCRDKMEVAFV 237
QY 241 CALNPSIDLHTSLELLQLOQLLWLLDGLHLYRPMALGNLADLEELEPTGPRDPLT 300
Db 238 CALNPSIDLHTSLELLQLOQLLWLLDGLHLYRPMALGNLADLEELEPTGPRDPLT 297
QY 301 LYHKGASAKTYRDBHIYPYMYLAGYHCRNRNRVREALQAWADATVTDYNYCREDEE 360
Db 298 IHLKAVESAKTYRDBHIYPYMYLAGYHCRNRNRVREALQAWADATVTDYNYCREDEE 357
QY 361 YKEFFVANDVPNLKKEAASLLEA--GEERPEQSGTOSQ--SALQDPECFAHLRLFYD 418
Db 358 YKEFFVANDVPNLKKEAASLLEA--GEERPEQSGTOSQ--SALQDPECFAHLRLFYD 417
QY 419 GICKWEESGTPVLHVGWATFVOSLGRFEGOVROKRVIVSREAEAAAEPEWGEAREG 478
Db 418 GICKWEESGTPVLHVGWATFVOSLGRFEGOVROKRVIVSREAEAAAEPEWGEAREG 477
QY 479 RRRGPRRESKPEEP-----PPKKPA-----LDKGLGTGGVAGVPPRPPGTVAGTARG 528
Db 478 RRRGPRRESKPEEP-----PPKKPA-----LDKGLGTGGVAGVPPRPPGTVAGTARG 528
QY 529 P-EGGSTAQVPAPASPPPE--GPVLTQSEKMKMKELLVANKINSATKLQTLTAOSQV 586
Db 534 KNEPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSPSP 593
QY 587 MKQKQVSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTP 610
Db 594 MKQKQVSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTPSTP 617

RESULT 4
Q9UE24 PRELIMINARY; PRT; 341 AA.
AC Q9UE24:
DT 01-MAY-2000 (Tremblrel. 13, Created)
DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
DE MRNA 3'-REGION (UNKNOWN FUNCTION) (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
ON NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98001089; PubMed=9341881;
RA Kedra D., Seroussi E., Fransson I., Trifunovic J., Clark M.,
RA Lagercrantz J., Blennow E., Menlin H., Dumanski J.;
RT "The germinal centre kinase gene and a novel CDC25-like gene are
RT located in the vicinity of the PYGM gene on 11q13."
RL Hum. Genet. 100:611-619(1997).
DR EMBL; Y12338; CAA73007.1; -.
FT NON_TER
SQ SEQUENCE 341 AA; 37723 MW; 01108F638C4CC64C CRC64;
```

Query Match 56.5%; Score 1800; DB 4; Length 341;
Best Local Similarity 100.0%; Pred. No. 6.8e-134;
Matches 341; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 270 LGHLERYPMALGNLADLEEPTGPRDPLTLYHKGIASAKTYRDEHIYPVYMLAGYHC 329

```
Db 1 LGHLERYPMALGNLADLEEPTGPRDPLTLYHKGIASAKTYRDEHIYPVYMLAGYHC 60
QY 330 RNRNRREALQAWADATVTDYNYCREDEEIIYKEFEVANDVTPNLKKEAASLLEAGEER 389
Db 61 RNRNRREALQAWADATVTDYNYCREDEEIIYKEFEVANDVTPNLKKEAASLLEAGEER 120
QY 390 PGEQSGTOSQSALQDPECFAHLRFYDGIKWEESGTPVLHVGWATFVOSLGRFEG 449
Db 121 PGEQSGTOSQSALQDPECFAHLRFYDGIKWEESGTPVLHVGWATFVOSLGRFEG 180
QY 450 QVRQKRVIVSREAEAAAEPEWGEAREGRRGRPRRESKPEEPKKPALDKGLGTGOG 509
Db 181 QVRQKRVIVSREAEAAAEPEWGEAREGRRGRPRRESKPEEPKKPALDKGLGTGOG 240
QY 510 AVSGPPRPPGTVAGTARGPEGSTAQVPAPASPPPEGVLTFQSEKMKMKELLVATK 569
Db 241 AVSGPPRPPGTVAGTARGPEGSTAQVPAPASPPPEGVLTFQSEKMKMKELLVATK 300
QY 570 INSSAIKLQTLTAOSQVOMKKQKVSTPSTPSTPSTPSTPSTPSTPSTPSTPSTP 610
Db 301 INSSAIKLQTLTAOSQVOMKKQKVSTPSTPSTPSTPSTPSTPSTPSTPSTPSTP 341

RESULT 5
Q9NKX0 PRELIMINARY; PRT; 751 AA.
AC Q9NKX0:
DT 01-OCT-2000 (Tremblrel. 15, Created)
DT 01-OCT-2000 (Tremblrel. 15, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE MENIN.
GN MN1 OR MEN1 OR CG13778.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
ON NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RA Maruyama K., Tsukada T., Honda M., Ohkura N., Nagasaki K., Kishi M.,
RA Nara-ashizawa N., Yamaguchi K.;
RT "Drosophila menin cDNA.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB040816; BAA94341.1; -.
DR FlyBase; FBgn0031885; Men1.
SQ SEQUENCE 751 AA; 81674 MW; D18EA61238C14115 CRC64;
```

Query Match 36.2%; Score 1153.5; DB 5; Length 751;
Best Local Similarity 38.1%; Pred. No. 2.1e-82;
Matches 272; Conservative 91; Mismatches 195; Indels 155; Gaps 15;

```
QY 9 TLFPLRSIDVVRLFAEL--GREEDLVLLSLVGVFHFVAVNRVPTNVPBLTQPS 67
Db 10 SLFPLKSTADVNLFRALTSIGTEPDLTLLSIWVGVIETSLTGTGAAAAQ-----AAQAAA 64
QY 68 ADPPPGGLTY-----FPVADLSITAIYARFTAQIRCAVDLSLYPREGV 112
Db 65 AAVAAGDISOATGGNDIIMGNSVPPFVTHETIAGLYKKFOT-IISVVE---KPKPHRQ 120
QY 113 SSRELKVKYSDVIWNSLSRSYFKDRAHIQSLFSGITGTCKLDSSGAVAFVVGACQALGLRD 172
Db 121 ATREVIKKYSDVIWNSLSRSYFKDRAHLQNLKLYSLGSKLDCFGVALAVVAGCOLLYK 180
QY 173 VHLALSEDHAWVVGNGEQTAEVTHWKGNEDERRGQTVNAGVAERSWLYLKGSMCRDR 232
Db 181 VRLAISEDHAWVVGNGEQTAEVTHWKGNEDERRGQTVNAGVAERSWLYLKGSMCRDR 240
QY 233 KMEYAVWCAINPSIDLHTSLELLQLOQLLWLLDGLHLYRPMALGNLADLEEPT 292
Db 241 GMEVAAICAAALNLSLTSNSDCVEAEVLEQQQLLWLLDGLHLYRPMALGNLADLEEPT 300
```

QY 293 PGRDPDLTYLHKGIASAKTYRDEHYIYPMYLAGYHCRNRNVRALQAMADTATVIQDYN 352
 Db 301 HPSTSCBOLYREATESARTHYRNHHVYPTYQNGYNNRLKYRDAFAAANAADVIRLYT 360
 QY 353 Y-CREDEEIKYKEFEVANDVIPNLLKEAASLEAGEERPEQSGTOSQO-----SALODP 407
 Db 361 YQCRDDEEIKYKELLDIANELIPYMK-----TESSGHSAIRLDS 401
 QY 408 ECFALHLLRFDYGICKWKEGSPPTVLHVGMATFLVQSLGRPEGOVROKVRT-----457
 Db 402 EVFANLLRFDYGICQWEDSLTPILHIGWAKPLVNNITKFDYDIRSOVVKLPEDLEAEQ 461
 QY 458 -----VSREA-EAAAEAPWGEAREGRRRR-----GPRRESKPEPP-----493
 Db 462 AKAEARAEQEAKEAKESKEAAGSEAMEGNNRNMATKEERNVLALCKELSSKPRNSOSSL 521
 QY 494 -----PPKKPALDKCL-----504
 Db 522 LLINISVLTETTSYLLAVLFTMLVHGLVSVFAESAMVLYKSKNSLPTWHLADLTAAAG 581
 QY 505 -----GTGQ-----GAVSGPP-----RKPPCTV 522
 Db 582 EKILNPDLFQGGQPFADQKQPSGGSNDPELHNNNSNNNNNNNNNNNNNNNNNNNNNNNN 641
 QY 523 AGTARGEGGSTAQVPAPASPPPEG-----548
 Db 642 TTNATTTNSGSGTSVQLPVSSSEANNAQAOQVQINDOLGKPOHKEAKKETSDYDPFE 701
 QY 549 -----PVLTFQSEKMGKMKELLVATKINSALQTAQSQVOMKOKVSTPS 596
 Db 640 IMLKRPVITLYSQMKGLKLLAEKLNTHAISLQVTAQSVASRKVRGVEKHS 692
 RESULT 6
 QYV47 PRELIMINARY; PRT; 813 AA.
 AC Q9VM47;
 DT 01-MAY-2000 (T-EMBLrel. 13, Created)
 DT 01-MAY-2000 (T-EMBLrel. 13, Last sequence update)
 DT 01-MAR-2001 (T-EMBLrel. 16, Last annotation update)
 DE CG13778 PROTEIN.
 GN MN1 OR CG13778.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Prorygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY.
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Vandal M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abrial J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berham B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam M.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista A.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levinsohn A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,

RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun Y.,
 RA Svirkas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster.";
 RL Science 287:2185-2195(2000).
 DR EMBL; AE003616; AAF52480.1; -.
 DR FlyBase; FBgn0031885; Mn1.
 DR InterPro; IPR001609; -.
 DR ProDom; PD000355; -; 1.
 SQ SEQUENCE 813 AA; 88384 MW; 4B7ED54B4BLC2AAF CRC64;

Query Match 35.3%; Score 1124.5; DB 5; Length 813;
 Best Local Similarity 35.4%; Pred. No. 4.5e-80;
 Matches 274; Conservative 88; Mismatches 198; Indels 213; Gaps 16;

QY 9 TLPLRSIDVRLFAEL-GREEPDLVLSVLGVFVHEFLAVNRVPTNVPELTQPSP 67
 Db 10 SLFPLKSTADVINFRRALTSGIEPDLTLLSVVGYELSLTTGEAAQ-----AAQAAA 64
 QY 68 APDPPGGLTY-----FPVADLSIIAALYARFTAQIRGAVDLSLYPREGGV 112
 Db 65 AAVAAGDISQATGGNDIIMGNVPPVPTVTHETAGLYKKFQT-ILSVB-----KPKPHRQ 120
 QY 113 SSRELKVKVSDVIWNSLSRYFKDRAHQISLFSFITGKLDSSGVAFVAVGCAQALGRD 172
 Db 121 ATREVKKVSDVIWNSLSIRSSYKDRALHQLNLYSLGKLDCEGVALAVVAGCOLLYGCR 180
 QY 173 VHLALEDHAWYVFGNGEQTAEVTHHGKCNEDRRGOTVAGVAERSWLYLKSGSYMCDR 232
 Db 181 VRLAISEDHAWYVFGKRVETIEVTHGKSGDKRGDIRPGEISGWSLYLGGLVAVVCR 240
 QY 233 KMEVAFMVCAINPSIDLHTLSLELQLOQLLWLLYDLGLHLYRYPMALGNLDELELEPT 292
 Db 241 GMEVAICAALNLSLTSNSDCVAELOQLLWLLYDLGLHLYRYPMALGTYLGELEIHR 300
 QY 293 PGRDPDLTYLHKGIASAKTYRDEHYIYPMYLAGYHCRNRNVRALQAMADTATVIQDYN 352
 Db 301 HPSTSCBOLYREATESARTHYRNHHVYPTYQNGYNNRLKYRDAFAAANAADVIRLYT 360
 QY 353 Y-CREDEEIKYKEFEVANDVIPNLLKEAASLEAGEERPEQSGTOSQO-----SALODP 407
 Db 361 YQCRDDEEIKYKELLDIANELIPYMK-----TESSGHSAIRLDS 401
 QY 408 ECFALHLLRFDYGICKWKEGSPPTVLHVGMATFLVQSLGRPEGOVROKVRT-----457
 Db 402 EVFANLLRFDYGICQWEDSLTPILHIGWAKPLVNNITKFDYDIRSOVVKLPEDLEAEQ 461
 QY 458 -----VSREA-EAAAEAPWGEAREGRRRR-----GPRRESKPEPP-----493
 Db 462 AKAEARAEQEAKEAKESKEAAGSEAMEGNNRNMATKEERNVLALCKELSSKPRNSOSSL 521
 QY 494 -----PPKKPALDKCL-----504
 Db 522 LLINISVLTETTSYLLAVLFTMLVHGLVSVFAESAMVLYKSKNSLPTWHLADLTAAAG 581
 QY 505 -----GTGQ-----GAVSGPP-----RKPPCTV 522
 Db 582 EKILNPDLFQGGQPFADQKQPSGGSNDPELHNNNSNNNNNNNNNNNNNNNNNNNNNNNN 641
 QY 523 AGTARGEGGSTAQVPAPASPPPEG-----548
 Db 642 TTNATTTNSGSGTSVQLPVSSSEANNAQAOQVQINDOLGKPOHKEAKKETSDYDPFE 701
 QY 549 -----PVLTFQSEKMGKMKELLVATKINSALQTAQSQVOMKOKVSTPS 596
 Db 640 IMLKRPVITLYSQMKGLKLLAEKLNTHAISLQVTAQSVASRKVRGVEKHS 692

Ds 702 IMKRPVITLYSQMKGLKDLLLAELKLNTHAISLQVTAQSVASRKVRGVKHS 754

RESULT 7

Q9CZQ5 ID Q9GZQ5 PRELIMINARY; PRT; 146 AA.
AC Q9GZQ5;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE MENIN (FRAGMENT).
GN MENI.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20484185; PubMed=11027505;
RA Khodaei-O'Brien S., Zabłowska B., Fromaget M., Bylund L., Weber G.,
RA Gaudray P.;
RL Biochem. Biophys. Res. Commun. 276:508-514(2000).
DR EMBL; AJ297489; CAC14133.1; -;
DR EMBL; AJ297485; CAC14129.1; -;
DR EMBL; AJ297486; CAC14130.1; -;
DR EMBL; AJ297487; CAC14131.1; -;
DR EMBL; AJ297488; CAC14132.1; -;
FT NON_TER 146
SQ SEQUENCE 146 AA; 16145 MW; BD7EDFF9A7AE03C7 CRC64;

Query Match 23.0%; Score 734; DB 4; Length 146;
Best Local Similarity 100.0%; Pred. No. 2.3e-50;
Matches 146; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MGLKAAQKTLFPLRSIDDDVRLFAELGREPDVLVLSVLGVFVHFVAVNRVPTNPVE 60
Db 1 MGLKAAQKTLFPLRSIDDDVRLFAELGREPDVLVLSVLGVFVHFVAVNRVPTNPVE 60
QY 61 LTFQSPAPDPGGTLTFYFVADLSIIAALYARFTAQIRGAVDSLTPREGVSSRELVRK 120
Db 61 LTFQSPAPDPGGTLTFYFVADLSIIAALYARFTAQIRGAVDSLTPREGVSSRELVRK 120
QY 121 VSDVWNSLSRSYFKDRAHIQSLFSF 146
Db 121 VSDVWNSLSRSYFKDRAHIQSLFSF 146

RESULT 8

Q24426 ID Q24426 PRELIMINARY; PRT; 531 AA.
AC Q24426;
DT 01-NOV-1996 (Tremblrel. 01, Created)
DT 01-MAY-1999 (Tremblrel. 10, Last sequence update)
DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
DE TROPOMYOSIN ISOFORM 33.
GN TM1 OR CG4898.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=84205681; PubMed=6202423;
RA Karik C.C., Mahaffey J.W., Coutu M.D., Fyrberg E.A.;
RT "Organization of contractile protein genes within the 88F subdivision
of the D. melanogaster third chromosome.";
RL Cell 37:469-481(1984).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87064486; PubMed=3097506;

RA Karik C.C., Fyrberg E.A.;
RT "Two Drosophila melanogaster tropomyosin genes: structural and
functional aspects.";
RL Mol. Cell Biol. 6:1965-1973(1986).
DR EMBL; K02621; AAA28968.1; -;
DR EMBL; M12840; AAA28968.1; JOINED.
DR EMBL; L00355; AAA28968.1; JOINED.
DR EMBL; L00356; AAA28968.1; JOINED.
DR EMBL; L00357; AAA28968.1; JOINED.
DR EMBL; L00358; AAA28968.1; JOINED.
DR EMBL; L00359; AAA28968.1; JOINED.
DR EMBL; L00360; AAA28968.1; JOINED.
DR EMBL; L00362; AAA28968.1; JOINED.
DR HSSP; P04002; 1ATF.
DR FlyBase; FBgn0003721; Tm1.
DR InterPro; IPR000104; -;
DR InterPro; IPR000533; -;
DR Pfam; PF00261; Tropomyosin; 2.
DR PRINTS; PR00308; ANTIFREEZEI.
DR PRINTS; PR01217; PRICHEXTENS.
DR PRINTS; PR00194; TROPOMYOSIN.
DR PROSITE; PS00326; TROPOMYOSIN; 1.
KW Alternative splicing.
SQ SEQUENCE 531 AA; 56047 MW; 2618A715E20EE4DF CRC64;

Query Match 4.4%; Score 139; DB 5; Length 531;
Best Local Similarity 21.3%; Pred. No. 0.013;
Matches 105; Conservative 50; Mismatches 194; Indels 144; Gaps 20;
QY 114 SRELKVKSDVIWNSLSRSYFKDRAHIQSLFSFITGKTLDSSGVAF-----AVVGACQAL 168
Db 43 ARQLQKKI-QTVENELDQI-----QEAALTIVTG-KLEEKNKALQNAESEVAALNRRI 92
QY 169 GLRDVHLALSEDHAWVFGPNGBQTAETVHGKGNDRGQTVNAGVARSWLYLKGSYM 228
Db 93 QLLEEDLERSEERSASAI---QLAAEASQSADESERARKILENRALADEERMDALENQL 148
QY 229 RCDRMEVAFWCAINPSIDLHTDLSLELLQLOOLKLLWLLYDLGLHLYPMALGNLADLEE 288
Db 149 K-----EARFLAEEADKKYD-----EVARKLAWVEADLERAEERAMV-----EADLER 191
QY 289 LEPTGCRPDPLTYLHKGIASAKTYRDEHYIPYMYLAGYHCRNRNVREALQAWADTATVI 348
Db 192 AEERAEQ-----GENKIVELEEE-----LRLVGNLKSLEYSE----- 224
QY 349 QDYNVCREDEEYKEFFEYVANDVIPNLKKEAASLLLEAGEEPGEQSQGTQSQGSALQD-- 406
Db 225 ---EKANQREEEYKQIKTLN----TRLKEA-----EAAEFAERSVOKLQKEVDRLDDDL 273
QY 407 -----PECFALLRFYDGICKWEGS---PTPVLHVGNWATFVLSLSGRFEG 449
Db 274 IVEKERYCMIGDSLDEAFVDLTKGLEPF--WNPKNPKPTPKL-----PTPTPEELAAEME 337
QY 450 QVRQKVRIVSRAEAAAEAEPEWGEAREGRRRGRPRRESKP-EPPEPPKPKALDKGLGTGQ 508
Db 328 ARAAEAAAAAEAAEAAAAAGACADGAPAPGEGAKAPEPTPKPEP----- 377
QY 509 GAVSGPPRPPGTPGTVAGTARGEGGTAQV-----PTTPEELAAEME 537
Db 378 ---TPPPPPPPPEYSIDLPEGAEPYVKNVEPPPPGSEPEVPAEAGEAFAEAGAAP 434
QY 538 PAPAASPPPEGPV 550
Db 435 PAEGAAPPAEGAV 447
RESULT 9
Q9VF97 ID Q9VF97 PRELIMINARY; PRT; 566 AA.
AC Q9VF97;
DT 01-MAY-2000 (Tremblrel. 13, Created)

DT 01-MAY-2000 (Tremblrel. 13, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE TMI OR CG4898.
 GN Drosophila melanogaster (Fruit fly).
 OS Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RX MEDLINE=20196006; PubMed=107311132;
 RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton R.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazer V.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattel B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wasman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR HSP: P04002; 1ATF.
 DR EMBL: AE003708; AAF5164.1; -.
 DR FlyBase: FBgn003721; Tm1.
 DR InterPro: IPR00104; -.
 DR InterPro: IPR000533; -.
 DR Pfam: PF00261; Tropomyosin; 2.
 DR PRINTS: PR00308; ANTIFREEZE1.
 DR PRINTS: PR01217; PRICHTEXTENS.
 DR PRINTS: PR00194; TROPOMYOSIN.
 DR PROSITE: PS00326; TROPOMYOSIN; 1.
 SQ SEQUENCE 566 AA; 60449 MW; 2AF07CFB0DA2032E CRC64;

Query Match 4.18; Score 129; DB 5; Length 566;
 Best Local Similarity 25.48; Pred. No. 0.086;
 Matches 62; Conservative 14; Mismatches 90; Indels 78; Gaps 9;

QY 358 EEIYKEFEFVANDVIPNLLKEAASLLEAGEERPGESQSGTOSGALQD----- 406
 DB 266 EEEYKNQIKTLN----FTLKEA----EAAFEAERSVQKQKQVDRLEDDLLIVEKERYCM 317

QY 407 -----PECFALLRFRYDGIKWEES---PTVLHVGMATFLVQSLGRFCQVRKQKRVIV 458
 DB 318 IGSLSDEAFVDLIKLEPF--WNRNPKPPTPKL---PTTPEELAAEMEAARAAEAAA 371
 QY 459 SRRAEAAEAEPEWGEAREGRRRGRPRRESKP-BEPPPPKPKPALDKGLCTCGGAVSGPPRK 517
 DB 372 AARAEAAEAAEAGEAGPDGAPAPAGEKAPAKEPTPKPE-----TPPPPP 418
 QY 518 PPGTVAGTARGPGGSTAQV-----PAPAAASPPP 546
 DB 419 PPPEYSIDLPPGEAEVYVKNYEPPEPPGSEPPVPAEAGEAAPAECAAPPA 478
 QY 547 EGPV 550
 DB 479 EGAV 482
 RESULT 10
 ID 013027 PRELIMINARY; PRT; 609 AA.
 AC 013027;
 DT 01-JUL-1997 (Tremblrel. 04, Created)
 DT 01-JUL-1997 (Tremblrel. 04, Last sequence update)
 DT 01-MAR-2001 (Tremblrel. 16, Last annotation update)
 DE MITOTIC PHOSPHOPROTEIN 90.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Piploidea; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97272720; PubMed=9115395;
 RA Stukenberg P.T., Lustig K.D., McGarry T.J., King R.W., Kuang J.,
 RA Kirschner M.W.;
 RT "Systematic identification of mitotic phosphoproteins."
 RL Curr. Biol. 7:338-348(1997).
 DR EMBL: U95102; AAC60123.1; -.
 DR InterPro: IPR001026; -.
 DR Pfam: PF01417; ENTH; 1.
 DR SMART: SM00273; ENTH; 1.
 SQ SEQUENCE 609 AA; 64589 MW; E65EC0CC954D2459 CRC64;
 Query Match 4.0%; Score 128.5; DB 13; Length 609;
 Best Local Similarity 23.7%; Pred. No. 0.1;
 Matches 71; Conservative 29; Mismatches 90; Indels 109; Gaps 14;
 QY 317 HIYPYMYLAGHCHNRNRVREALQAWADTAIV--IQDYNCEDEEIEYKEFEFVANDVIPN 374
 DB 64 HVKAMTLMETIKTGSERVAQCKENIATQTLKQFYVDRO-----GKQGVN 113
 QY 375 LLKEAASLLEAGEERPGESQSGTOSGALQDPECF---AHLRFYDGIKWEES--- 427
 DB 114 VREKAKOLV-----SLKDDERKEERAAHALKTKREK(AOT)STSSAS 155
 QY 428 ---PTVLHVGMATFLVQSLGRFGQVRKRVISREAEAAEAEPP----- 470
 DB 156 STLNPAPEGEQAWS---QSSGEELQLALAMSKEAEQVRAKPPVSEELQLAL 211
 QY 471 -----WGEEAR-----EGRRGPRRESKPEE-----PPPKPKP 498
 DB 212 SLSKEEHDKKEERIKRGDDLRLQMALESRRKGP---SKQEEQSLSMDLADVFSPPAPVAP 268
 QY 499 ALDKGLCTGCGAVSGPPKPKPGTAVGTARGEGGSTAQVPAPAPASPPPEG--PVLTFQS 555
 DB 269 TADP-----WGASAPPPADP---WAG-----GATPASVPANAAPDPWGGPPVATGSS 313
 RESULT 11
 ID 075229 PRELIMINARY; PRT; 813 AA.
 AC 075229;


```
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE BICAUDAL D HOMOLOG 1 (FRAGMENT).
GN BICD1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96079090; PubMed=8530100;
RA Baens M., Aerssens J., van Zand K., Van Den Berghe H., Marynen P.;
RT "Isolation and regional assignment of human chromosome 12p cDNAs.";
RL Genomics 29:44-52(1995).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98035884; PubMed=9367685;
RA Baens M., Marynen P.;
RT "A human homologue (BICD1) of the Drosophila bicaudal-D gene.";
RL Genomics 45:601-606(1997).
DR EMBL; U90029; AAB94807.1; -.
DR MGD; MGI:1101760; Bicd1.
FT NON_TER 1
FT NON_TER 545
SQ SEQUENCE 545 AA; 61962 MW; D4EB6DD6748C5F6 CRC64;

Query Match 4.0%; Score 126; DB 11; Length 545;
Best Local Similarity 19.3%; Pred. No. 0.14;
Matches 94; Conservative 75; Mismatches 175; Indels 142; Gaps 21;

Qy 166 QALGLRDVHLALSEDHAWVFPNGEQTAEVTHGKGNDRGQTVNAGVAERSWLYLK 225
Db 1 QINLSDSHISVSD--GLKAEADGSE-----PNNDDKMNGHIHGLPKG-----LNG 68

Qy 226 SYMR-CDRMEVAFWVCAINSIDLHTD--SLELLQLOQKLLWLLYDLGLHLYRPMALGN 282
Db 69 DYRTPTTRKGE-----SLHPVSLFSELNISEIOLKQQLI-----QVEREKAIL-- 113

Qy 283 LADLEE----LEPTPGRPDPLTLVH-----KGIAKATVYRD----- 315
Db 114 LANQESQTQLEHTKG---ALTEQHERVHRLTEHVNAMRGLQNSKEIKAEILDCRGRNSA 170

Qy 316 EHVYPYMY----LAGYHGRNR-NVREALQAWADTATVTDQNYNC-----REDEIYKEFFE 366
Db 171 BEAHDYEVDINGLEILECKYRVAVTEVIDLKAELKALKEKYNKSNYENYTEKTKYESKIQ 230

Qy 367 VANDVIPNLLKASLLLEAGE-----ERPGEOSQGTOSGSALODPECEFA-HLLR 415
Db 231 MYDEQVTLNKKTSK---ESGEKMXHMEKELOKMTGIANENHNTLNTAODELVTFSEELAQ 287

Qy 416 FYDGICKWEEGSPTPVLHVGWATFLVQSLGRFEG-----QVRQKRVIVSREAEAAEA 467
Db 288 LYHHVCLCNNETPNRMVLDIYRQSVTRSGSLKGFDDPRGLLSRLSRGVSSPVESRTS 347

Qy 468 EEPNGEAEARRRRPRRESKPEPPPPKPKALDKGLGTGOGAVSGPPPKPGTVACTAR 527
Db 348 SEPVSKEKTE-----TSKPSPTKTPTI----- 370

Qy 528 GPEGSTAQVAPAPASPPPEGPVLTFQSEKMKMKELLVATKINSIAIKLQLTAQSQVQM 587
Db 371 -----SPVITAPPSPVLDTSIRKEPNMNYNLNAILRDIQTKHLQKAVDRSLQL 419

Qy 588 KQKQVS 593
Db 420 SRQRAA 425

RESULT 14
Q9JLT1 PRELIMINARY; PRT; 4880 AA.
ID Q9JLT1
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
```

```
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE MULTIDOMAIN PRESYNAPTIC CYTOMATRIX PROTEIN PICCOLO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20170257; PubMed=10707984;
RA Fenster S.D., Chung W.J., Zhai R., Cases-Langhoff C., Voss B.,
RA Garner A.M., Kaempf U., Kindler S., Gundelfinger E.D., Garner C.C.;
RT "Piccolo, a Presynaptic Zinc Finger Protein Structurally Related to
RT Bassoon.";
RL Neuron 25:203-214(2000).
DR EMBL; AF138789; AAF07822.2; -.
DR InterPro; IPR000008; -.
DR InterPro; IPR001478; -.
DR Pfam; PF001168; C2; 1.
DR PRINTS; PR01217; PRICHTEXTENS.
DR PROSITE; PS00499; C2_DOMAIN_1; UNKNOWN_1.
DR PROSITE; PS50004; C2_DOMAIN_2; 1.
DR SMART; SM00239; C2; 1.
KW Matrix protein.
SQ SEQUENCE 4880 AA; 530148 MW; 73951EE4ED83EA68 CRC64;

Query Match 3.9%; Score 125.5; DB 11; Length 4880;
Best Local Similarity 22.0%; Pred. No. 3.8;
Matches 69; Conservative 47; Mismatches 80; Indels 117; Gaps 18;

Qy 291 PTPGRPDPLTLVHKGIAASAKT--YYRDEHIYPYMYLAGYHCRNRNRVREALQAWADTATVI 348
Db 96 PDGCRPT-----QFGLSKSRITDTFRSEQKLP-----GRSPSTIS 130

Qy 349 QQYNYCRED-EETIYKE-----FF-----EVANDVIP-NLLKRA-ASLLAEGERP 390
Db 131 LKESKRTDFKEEYKSSMMPGFFSDVNPVLSAVSSVWKNFNPFDLISDSEASQEBETTKQK 190

Qy 391 GEOSQGTOSGSA---LQDPECFALLRFFYDGICKWEEGSPTPVLHVGWATFLVQSLGRF 447
Db 191 VVQKEQKSGEAKPPLOQP-----SPKPI-----PKQ 218

Qy 448 EGQVRQKVR-----IVSREAEAAEAEEP-WGEAREGRRGRPRRESKP----- 489
Db 219 QGQYKEVIQDSSPKSVSSQQAQKVPQAGTGPQSPQAQTPAQQASPKVPVQAQPGS 278

Qy 490 -----EPPPPPKPKALDKGLGTGOGAVSGPPPKPGTVAGTARGEGGSTAQVPAP---A 541
Db 279 AKATVQQPGPAKSPA--QPAGTGKS-----PAQPPAKTPGQQAAGLEKTSSTSSQQPGPKSLA 331

Qy 542 ASPP-----PEGPV 550
Db 332 QTPGHGKFPPLGPV 344

RESULT 15
Q9JKS6 PRELIMINARY; PRT; 5085 AA.
ID Q9JKS6
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE MULTIDOMAIN PRESYNAPTIC CYTOMATRIX PROTEIN PICCOLO.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RX Fenster S.D., Cases-Langhoff C., Gundelfinger E.D., Garner C.C.;
RT "Multidomain presynaptic cytomatrix protein Piccolo, long splice
```

no submitted (from 2000) to the EMBL/GenBank/DDBJ databases

DR	EMBL; AF227534; AAF63196.1; -
DR	InterPro; IPR000008; -
DR	InterPro; IPR001478; -
DR	InterPro; IPR002965; -
DR	Pfam; PF00168; C2; 2.
DR	PRINTS; PR01217; PRICHTEXTENS.
DR	PROSITE; PS00499; C2.DOMAIN.1;
DR	PROSITE; PS00004; C2.DOMAIN.2; 2.
DR	SMART; SM00239; C2; 1.
KW	Matrix protein.
SQ	SEQUENCE 5085 AA; 552711 MW; 5A1BB543201A7450 CRC64;

Search completed: October 12, 2001, 18:21:53
Job time: 201 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:18:52 ; Search time 13.19 seconds
(without alignments)
1584.219 Million cell updates/sec

Title: US-09-380-337-2

Perfect score: 3185

Sequence: 1 MGLKAAQKTLFLRSIDV.....KVSTPSDYTLSELRQRKGL 610

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 93435 seqs, 34255486 residues

Total number of hits satisfying chosen parameters: 93435

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_39.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3172.5	99.6	615	1 MEN1_HUMAN	O00255 homo sapien
2	3049.5	95.7	611	1 MEN1_MOUSE	O88359 mus musculus
3	141.5	4.4	518	1 TPM4_DROME	P49455 drosophila
4	132	4.1	1516	1 CALH_HUMAN	P39060 homo sapien
5	125.5	3.9	1025	1 CAL6_MOUSE	O04857 mus musculus
6	124.5	3.9	992	1 EBN6_EBV	P03204 epstein-bar
7	123	3.9	331	1 PRP1_HUMAN	P04280 homo sapien
8	122.5	3.8	1790	1 SEPA_EMENI	P78621 emericella
9	119	3.7	172	1 PRP2_RAT	P10164 rattus norv
10	118	3.7	705	1 PRP2_HUMAN	P17600 homo sapien
11	117.5	3.7	251	1 PRP2_MOUSE	P02812 homo sapien
12	117.5	3.7	1581	1 PRPB_HUMAN	O15648 h peroxisom
13	117	3.7	308	1 MACS_MOUSE	P26645 mus musculus
14	116	3.6	704	1 SYNL_RAT	P09951 rattus norv
15	116	3.6	726	1 ADDB_HUMAN	P35612 homo sapien
16	116	3.6	860	1 AREA_PENRO	O13508 penicillium
17	116	3.6	2774	1 MAPA_RAT	P34926 rattus norv
18	115.5	3.6	247	1 PRP4_HUMAN	P10163 homo sapien
19	114	3.6	492	1 MEC2_RAT	O00566 rattus norv
20	113.5	3.6	479	1 ACMA_HUMAN	P08173 homo sapien
21	113.5	3.6	1838	1 CAL5_HUMAN	P20908 homo sapien
22	113	3.5	484	1 MEC2_MOUSE	O92246 mus musculus
23	113	3.5	706	1 SYNL_BOVIN	P17599 bos taurus
24	112.5	3.5	384	1 AIPL_HUMAN	O9n2n9 homo sapien
25	111.5	3.5	861	1 DYN1_MOUSE	P39053 mus musculus
26	111.5	3.5	2453	1 NCRL_MOUSE	O60974 mus musculus
27	111	3.5	276	1 PRPL_HUMAN	P10162 homo sapien
28	111	3.5	520	1 WASP_MOUSE	P70315 mus musculus
29	111	3.5	1888	1 CAL6_CHICK	P32018 gallus gall
30	110.5	3.5	549	1 AGLA_RHIME	O92318 rhizobium m
31	110	3.5	174	1 PRPP_HUMAN	P81489 homo sapien
32	110	3.5	479	1 ACMA_MOUSE	P32211 mus musculus
33	110	3.5	801	1 BRD2_HUMAN	P25440 homo sapien

RESULT	ID	MEN1_HUMAN	STANDARD;	PRT;	615 AA.
34	110	3.5	857	1 NFM_CHICK	P16053 gallus gall
35	110	3.5	2220	1 YICI_HUMAN	O9y610 homo sapien
36	109.5	3.4	772	1 YICI_ECOLI	P31434 escherichia
37	109	3.4	670	1 SYNL_MOUSE	O88935 mus musculus
38	109	3.4	783	1 FYB_HUMAN	O15117 homo sapien
39	108.5	3.4	864	1 DYN1_HUMAN	O05193 homo sapien
40	108.5	3.4	865	1 NRFA_PENUR	O92269 penicillium
41	108	3.4	342	1 ME18_MOUSE	P23798 mus musculus
42	108	3.4	689	1 L100_ADECC	O65957 canine aden
43	108	3.4	2205	1 POLN_RUBVT	P13889 rubella vir
44	107.5	3.4	384	1 VASP_CANFA	P50551 canis famil
45	106.5	3.3	308	1 MACS_RAT	P30009 rattus norv

ALIGNMENTS

RESULT	ID	MEN1_HUMAN	STANDARD;	PRT;	615 AA.
34	110	3.5	857	1 NFM_CHICK	P16053 gallus gall
35	110	3.5	2220	1 YICI_HUMAN	O9y610 homo sapien
36	109.5	3.4	772	1 YICI_ECOLI	P31434 escherichia
37	109	3.4	670	1 SYNL_MOUSE	O88935 mus musculus
38	109	3.4	783	1 FYB_HUMAN	O15117 homo sapien
39	108.5	3.4	864	1 DYN1_HUMAN	O05193 homo sapien
40	108.5	3.4	865	1 NRFA_PENUR	O92269 penicillium
41	108	3.4	342	1 ME18_MOUSE	P23798 mus musculus
42	108	3.4	689	1 L100_ADECC	O65957 canine aden
43	108	3.4	2205	1 POLN_RUBVT	P13889 rubella vir
44	107.5	3.4	384	1 VASP_CANFA	P50551 canis famil
45	106.5	3.3	308	1 MACS_RAT	P30009 rattus norv

RP VARIANTS FMEN1.
RX MEDLINE=98349969; PubMed=9683585;
RA Giraud S., Zhang C.X., Serova-Sinilnikova O., Wautot V., Salandre J.,
RA Bulisson N., Waterlot C., Bauders C., Porchet N., Aubert J.-P., Emy P.,
RA Cadot G., Delamer B., Chabre O., Niccoli P., Lepat F., Duron F.,
RA Emperauger B., Cougard P., Goudet P., Sarfati E., Riou J.-P.,
RA Guichard S., Rodier M., Meyrier A., Caron P., Vantghem M.-C.,
RA Assayag M., Peix J.-L., Pugeat M., Rohmer V., Vallotton M., Lenoir G.,
RA Gaudray D., Proye C., Conte-Devolx B., Chanson P., Shugart Y.Y.,
RA Goldgar D., Murat A., Calender A.;
RT "Germline mutation analysis in patients with multiple endocrine
RT neoplasia type 1 and related disorders.";
RL Am. J. Hum. Genet. 63:455-467(1998).
[6]
RP VARIANT FIHP LYS-260.
RX MEDLINE=99011276; PubMed=9792884;
RA Teh B.T., Esapa C.T., Houlston R., Grandell U., Farnebo F.,
RA Nordensjoeld M., Pearce C.J., Carmichael D., Larsson C., Harris P.E.;
RT "A family with isolated hyperparathyroidism segregating a missense
RT MEN1 mutation and showing loss of the wild-type alleles in the
RT parathyroid tumors.";
RL Am. J. Hum. Genet. 63:1544-1549(1998).
[7]
RP VARIANT FIHP GLU-189.
RX MEDLINE=99057176; PubMed=9843042;
RA Fujimori M., Shirahama S., Sakurai A., Hashizume K., Hama Y., Ito K.,
RA Shingu K., Kobayashi S., Anano J., Fukushima Y.;
RT "Novel V184E MEN1 germline mutation in a Japanese kindred with
RT familial hyperparathyroidism.";
RL Am. J. Med. Genet. 80:221-222(1998).
[8]
RP VARIANTS FMEN1.
RX MEDLINE=98334342; PubMed=9671267;
RA Agarwal S.K., Debelenko L.V., Kester M.B., Guru S.C., Manickam P.,
RA Olufeni S.-E., Skarulis M.C., Heppner C., Crabtree J.S.,
RA Lubensky I.A., Zhuang Z., Kim Y.S., Chandrasekharappa S.C.,
RA Collins F.S., Liotta L.A., Spiegel A.M., Burns A.L., Emmert-Buck M.R.,
RA Marx S.J.;
RT "Analysis of recurrent germline mutations in the MEN1 gene encountered
RT in apparently unrelated families.";
RL Hum. Mutat. 12:75-82(1998).
[9]
RP VARIANT FMEN1 ILE-135 AND LYS-364.
RX MEDLINE=98410971; PubMed=9740255;
RA Boeni R., Vormeyer A.O., Pack S., Park W.-S., Burg G., Hofbauer G.,
RA Darling T., Liotta L., Zhuang Z.;
RT "Somatic mutations of the MEN1 tumor suppressor gene detected in
RT sporadic angiofibromas.";
RL J. Invest. Dermatol. 111:539-540(1998).
[10]
RP VARIANTS FMEN1 LYS-119 DEL AND GLN-171--LEU-173 DEL.
RX MEDLINE=98419173; PubMed=9747036;
RA Sakurai A., Shirahama S., Fujimori M., Katai M., Itakura Y.,
RA Kobayashi S., Anano J., Fukushima Y., Hashizume K.;
RT "Novel MEN1 gene mutations in familial multiple endocrine neoplasia
RT type 1.";
RL J. Hum. Genet. 43:199-201(1998).
[11]
RP VARIANT FMEN1 GLY-45.
RX MEDLINE=99048878; PubMed=9832038;
RA Sato M., Matsubara S., Miyauchi A., Ohye H., Imachi H., Murao K.,
RA Takahara J.;
RT "Identification of five novel germline mutations of the MEN1 gene in
RT Japanese multiple endocrine neoplasia type 1 (MEN1) families.";
RL J. Med. Genet. 35:915-919(1998).
[12]
RP VARIANTS FMEN1 TRP-39; TYR-177; ASP-184; PRO-269 AND PRO-272.
RX MEDLINE=99103464; PubMed=9886389;
RA Poncin J., Abs R., Velkeniers B., Bonduelle M., Abranowicz M.,
RA Legros J.-J., Verloes A., Meurisse M., van Gaal L., Verellen C.,
RA Koulischer L., Beckers A.;
RT "Mutation analysis of the MEN1 gene in Belgian patients with multiple
RT endocrine neoplasia type 1 and related diseases.";

Hum. Mutat. 13:54-60(1999).
[13]
RP VARIANTS MEN1 ASP-161 AND ARG-246.
RX MEDLINE=99188881; PubMed=10090472;
RA Mutch M.G., Dilley W.G., Sanjurjo F., Debenedetti M.K., Doherty G.M.,
RA Wells S.A. Jr., Goodfellow P.J., Laimore T.C.;
RT "Germline mutations in the multiple endocrine neoplasia type 1 gene:
RT evidence for frequent splicing defects.";
RL Hum. Mutat. 13:175-185(1999).
CC -1- FUNCTION: NOT KNOWN.
CC -1- SUBCELLULAR LOCATION: NUCLEAR.
CC -1- TISSUE SPECIFICITY: UBIQUITOUS.
CC -1- DISEASE: DEFECTS IN MEN1 IS THE CAUSE OF FAMILIAL MULTIPLE
CC ENDOCRINE NEOPLASIA TYPE 1 (FMEN1); WERNER SYNDROME, AN AUTOSOMAL
CC DOMINANT DISORDER CHARACTERIZED BY TUMORS OF THE PARATHYROID
CC GLANDS, GASTRO-INTESTINAL ENDOCRINE TISSUE, THE ANTERIOR PITUITARY
CC AND OTHER TISSUES. CUTANEOUS LESIONS AND NERVOUS-TISSUE TUMORS
CC CAN EXIST. PROGNOSIS IN FMEN1 PATIENTS IS RELATED TO HORMONAL
CC HYPERSECRETION BY TUMORS, SUCH AS HYPERGASTRINEMIA CAUSING SEVERE
CC PEPTIC ULCER DISEASE (ZOLLINGER-ELLISON SYNDROME, ZES), PRIMARY
CC HYPERPARATHYROIDISM, AND ACUTE FORMS OF HYPERINSULINEMIA.
CC -1- DISEASE: DEFECTS IN MEN1 IS THE CAUSE OF FAMILIAL ISOLATED
CC HYPERPARATHYROIDISM (FIHP OR HRPT1). FIHP IS AN AUTOSOMAL DOMINANT
CC DISORDER CHARACTERIZED BY HYPERCALCEMIA, ELEVATED PARATHYROID
CC HORMONE (PTH) LEVELS, AND UNIGLANDULAR OR MULTIGLANDULAR
CC PARATHYROID TUMORS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U93236; AAC51228.1; -;
DR EMBL; U93237; AAC51229.1; -;
DR EMBL; U93237; AAC51230.1; -;
DR MIM; 131100; -;
DR MIM; 145000; -;
KW Nuclear protein; Disease mutation: Alternative splicing; Polymorphism.
FT VARSPLIC 149 153 MISSING (IN SHORT ISOFORM).
FT VARIANT 12 12 P-> L (IN FMEN1).
FT VARIANT 22 22 L-> R (IN FMEN1).
FT VARIANT 26 26 E-> K (IN PARATHYROID ADENOMA).
FT VARIANT 39 39 L-> W (IN FMEN1).
FT VARIANT 42 42 G-> D (IN FMEN1).
FT VARIANT 45 45 E-> G (IN FMEN1).
FT VARIANT 119 119 MISSING (IN FMEN1).
FT VARIANT 135 135 K-> I (IN FMEN1).
FT VARIANT 139 139 H-> D (IN FMEN1).
FT VARIANT 139 139 H-> Y (IN FMEN1 AND SPORADIC MEN1).
FT VARIANT 144 144 F-> V (IN FMEN1).
FT VARIANT 161 161 G-> D (IN FMEN1).
FT VARIANT 165 165 A-> P (IN FMEN1).
FT VARIANT 169 169 A-> D (IN FMEN1).
FT VARIANT 171 173 MISSING (IN FMEN1).
FT VARIANT 176 176 R-> Q.

SEQUENCE FROM N.A.									
RP	Query Match	99.6%	Score 3172.5;	DB 1;	Length 615;				
RX	Best Local Similarity	99.2%	Pred. No. 99-199;						
RA	Matches 610;	Conservative	0;	Mismatches	0;	Indels	5;	Gaps	1;
QY	1	MGLKAAQKTLFPLRSIDDDVRLFAAELGREEDLVLLSLVLGFEVHFHAVNRVPTNPVE	60						
DB	1	MGLKAAQKTLFPLRSIDDDVRLFAAELGREEDLVLLSLVLGFEVHFHAVNRVPTNPVE	60						
QY	61	LTFQSPADPPGGGLTYFPVADLSIIAALYARETAQIRGAVDLSLYPREGGVSSRELVKK	120						
DB	61	LTFQSPADPPGGGLTYFPVADLSIIAALYARETAQIRGAVDLSLYPREGGVSSRELVKK	120						
QY	121	VSDVIWNSLSRSYFKDRAHIQSLFSFITT-----GTKLDSGGVAFVAVGACQALGRDVLH	175						
DB	121	VSDVIWNSLSRSYFKDRAHIQSLFSFITT-----GTKLDSGGVAFVAVGACQALGRDVLH	180						
QY	176	ALSEDHAWVVGPNGBQTAEVTHWKGNEEDRRGQTVNAGVAERSWLYLKGSMRCDRKME	235						
DB	181	ALSEDHAWVVGPNGBQTAEVTHWKGNEEDRRGQTVNAGVAERSWLYLKGSMRCDRKME	240						
QY	236	VAFMVCAINPSIDLHTDLSLELLOQLLWLLYDLGLHLEYPMALGNLADLEELEPTGR	295						
DB	241	VAFMVCAINPSIDLHTDLSLELLOQLLWLLYDLGLHLEYPMALGNLADLEELEPTGR	300						
QY	296	POPLTYLHKGIAASAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQDYNCR	355						
DB	301	POPLTYLHKGIAASAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQDYNCR	360						
QY	356	EBEIIYKEFEFVANDVIPNLLKEAASLLBAGERPGEQSGQTSQSALQDPCEFAHLR	415						
DB	361	EBEIIYKEFEFVANDVIPNLLKEAASLLBAGERPGEQSGQTSQSALQDPCEFAHLR	420						
QY	416	FYDGICKWEESGTPVLHVGMATFLVQSLRGREGQVRQKRVIRSVREAAEAEPGGEA	475						
DB	421	FYDGICKWEESGTPVLHVGMATFLVQSLRGREGQVRQKRVIRSVREAAEAEPGGEA	480						
QY	476	REGRRGRPRRESKPEPPPPKPKALDKGLTGCGAVSGPPKPPGTVAGTARGPEGGSTA	535						
DB	481	REGRRGRPRRESKPEPPPPKPKALDKGLTGCGAVSGPPKPPGTVAGTARGPEGGSTA	540						
QY	536	QVPAPASPPPEGPVLTFQSEKMKMKELLVATKINSIAIKLQLTQSQVQMKKQKVSTP	595						
DB	541	QVPAPASPPPEGPVLTFQSEKMKMKELLVATKINSIAIKLQLTQSQVQMKKQKVSTP	600						
QY	596	SDYTLFLKRQRKGL 610							
DB	601	SDYTLFLKRQRKGL 615							
RESULT 2									
MEN1_MOUSE	STANDARD;	PRT;	611	AA.					
AC	O88559;								
DT	15-DEC-1998 (Rel. 37, Created)								
DT	15-DEC-1998 (Rel. 37, Last sequence update)								
DT	01-OCT-2000 (Rel. 40, Last annotation update)								
DE	MEN1.								
GN	MEN1.								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
OX	NCBI_TaxID=10090;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=129/OLA;								
RX	MEDLINE=9911257; PubMed=9893060;								
RA	Bassett J.H.D., Rashbass P., Harding B., Forbes S.A., Pannett A.A.,								
RA	Thakker R.V.;								
RT	"Studies of the murine homolog of the multiple endocrine neoplasia								
RT	type 1 (MEN1) gene, men1."								
RL	J. Bone Miner. Res. 14:3-10(1999).								
RN	[2]								

RA Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Soeda E.,
 RA Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K., Polley A.,
 RA Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
 RA Reichwald A., Rump A., Shillhabel M., Schudy A., Zimmermann W.,
 RA Rosenthal K., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
 RA Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
 RA Minoshima S., Shimizu N., Nordstiek G., Hornischer K., Brandt P.,
 RA Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloeker H.,
 RA Ramsperger S., Borym K., Klages S., Hennig S., Riesselmann L., Dagand E.,
 RA Wehrmeyer S., Borym K., Gardiner K., Nizetic D., Francis F.,
 RA Lehrach H., Reinhardt R., Yaspo M.-L.;
 RT "The DNA sequence of human chromosome 21.";
 RL Nature 405:311-319(2000).
 RN [3]
 RN SEQUENCE OF 834-1516 FROM N.A.
 RP MEDLINE=94245237; PubMed=8188291;
 RX Oh S.-P., Warman M.L., Seidlin M.F., Cheng S., Knoll J.H., Timmons S.,
 RA Olsen B.R.;
 RT "Cloning of cDNA and genomic DNA encoding human type XVIII collagen
 RT and localization of the alpha 1(XVIII) collagen gene to mouse
 RT chromosome 10 and human chromosome 21.";
 RL Genomics 19:494-499(1994).
 RN [4]
 RN SEQUENCE OF 1334-1516 FROM N.A.
 RP TISSUE=Placenta;
 RX Zhi-Yong H., Biao L., Wei-Jie Z., Xiang-Fu W.;
 RA "Cloning and expression of human endostatin gene in Escherichia
 RL coli.";
 RT Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RN INVOLVEMENT IN KNOBLOCH SYNDROME.
 RP MEDLINE=20400145; PubMed=10942434;
 RX Sertile A.L., Mossi V., Camargo A.A., Zatz M., Brahe C.,
 RA Passos-Bueno M.R.;
 RA "Collagen XVIII, containing an endogenous inhibitor of angiogenesis
 RT and tumor growth, plays a critical role in the maintenance of retinal
 RT structure and in neural tube closure.";
 RL Hum. Mol. Genet. 9:2051-2058(2000).
 CC -!- FUNCTION: COLA18A PROBABLY PLAYS A MAJOR ROLE IN DETERMINING THE
 CC RETINAL STRUCTURE AS WELL AS IN THE CLOSURE OF THE NEURAL TUBE.
 CC -!- FUNCTION: ENDOSTATIN POTENTIALLY INHIBITS ENDOTHELIAL CELL
 CC PROLIFERATION AND ANGIOGENESIS. MAY INHIBIT ANGIOGENESIS BY
 CC BINDING TO THE HEPARAN SULPHATE PROTEGLYCANS INVOLVED IN GROWTH
 CC FACTOR SIGNALLING (BY SIMILARITY).
 CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A SHORT FORM/NC1-303 AND A LONG
 CC FORM/NC-493 (SHOWN HERE); ARE PRODUCED BY ALTERNATIVE SPLICING.
 CC -!- TISSUE SPECIFICITY: PRESENT IN MULTIPLE ORGANS WITH HIGHEST LEVELS
 CC IN LIVER, LUNG AND KIDNEY.
 CC -!- PTM: PROLINES AT THE THIRD POSITION OF THE TRIPEPTIDE REPEATING
 CC UNIT (G-X-Y) ARE HYDROXYLATED IN SOME OR ALL OF THE CHAINS.
 CC -!- DISEASE: DEFECTS IN COL18A1 ARE A CAUSE OF KNOBLOCH SYNDROME (KS);
 CC AN AUTOSOMAL RECESSIVE DISORDER DEFINED BY THE OCCURRENCE OF HIGH
 CC MYOPIA, VITREORETINAL DEGENERATION WITH RETINAL DETACHMENT,
 CC MACULAR ABNORMALITIES AND OCCIPITAL ENCEPHALOCELE.
 CC -!- SIMILARITY: BELONGS TO THE MULTIPLEXIN FAMILY OF COLLAGENS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF018081; AAC39658.1; -;
 CC DR EMBL; AF018082; AAC39659.1; -;
 CC DR EMBL; AL163302; CAB90482.1; -;
 CC DR EMBL; L22548; AAA51864.1; -;
 CC DR EMBL; AF184060; AAF01310.1; ALT_INIT.
 CC DR GlycoSuiteDB; P39060; -;
 CC DR MIM; 120328; -;
 CC DR MIM; 267750; -;
 CC DR InterPro; IPR000087; -;

GN BERP3-BERF4.
OS Epstein-Barr virus (strain B95-8) (Human herpesvirus 4).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Gammaherpesvirinae; Lymphocryptovirus.
ON NCBI_TaxID=10377;
RX SEQUENCE FROM N.A.
RX MEDLINE=84270667; PubMed=6087149;
RA Baer R., Bankier A.T., Biggin M.D., Deininger P.L., Farrell P.J.,
RA Gibson T.J., Hatfull G., Hudson G.S., Satchwell S.C., Sequin C.,
RA Tuffnell P.S., Barrell B.G.;
RT "DNA sequence and expression of the B95-8 Epstein-Barr virus genome.";
RT Nature 310:207-211(1984).
RN [2]
RN CHARACTERIZATION.
RX MEDLINE=88155772; PubMed=2831394;
RX Petti L., Sample J., Wang F., Kieff E.;
RA "A fifth Epstein-Barr virus nuclear protein (EBNA3C) is expressed in
RT latently infected growth-transformed lymphocytes.";
RL J. Virol. 62:1330-1338(1988).
RN [3]
RN SUBCELLULAR LOCATION
RX MEDLINE=90266473; PubMed=2161150;
RA Petti L., Sample C., Kieff E.;
RT "Subnuclear localization and phosphorylation of Epstein-Barr virus
RT latent infection nuclear proteins.";
RL Virology 176:563-574(1990).
CC -!- FUNCTION: INVOLVED IN LATENT CYCLE.
CC -!- SUBCELLULAR LOCATION: NUCLEAR. ASSOCIATED WITH THE NUCLEAR MATRIX.
CC -!- SIMILARITY: SOME SIMILARITIES EXIST BETWEEN EBNA 4, 5, AND 6.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
DR EMBL; V01555; CAA24859.1; -
DR PIR; A03766; Q0BE25.
KW Nuclear protein; Repeat.
FT DOMAIN 74 80 POLY-ARG.
FT DOMAIN 551 610 10 X 5 AA TANDEM REPEATS.
FT DOMAIN 741 779 3 X 13 AA TANDEM REPEATS.
SQ SEQUENCE 992 AA; 109129 MW; 39BEAB9BC515BD84 CRC64;
Query Match 3.9%; Score 124.5; DB 1; Length 992;
Best Local Similarity 20.1%; Pred. No. 1.2;
Matches 112; Conservative 53; Mismatches 157; Indels 235; Gaps 25;
QY 137 RAHQSLFETIKLSDSSGVAFVAGCAQLGRDVH-----LALSEDH 181
DB 121 QSHLQALSNLIDSLDGLDTHILCFVMAARQR--LQDIRGLPLVAEGVGWRHLLTSPSQ 178
QY 182 AWVYFGPNGEQTAEV-----TWKGNEDRRGQTVNAGV----- 215
DB 179 SW---PMGYRTATLTLPVNRVGCADSIMLTATGCGNAQTLNTFTSVTPPHAGP 234
QY 216 -----AERSWLYKSGVRCRDKKEVAFVCAINPSIDLFTSLQLQOKLWLLY 268
DB 235 REQERYAREAEVFLRGKWRRIYDILFELCG-----SLHHIWNQLLQTEENLLDFVR 289
QY 269 DLGHLERYPMALGNLADLELEPTPGRPDPLTLYHKGIASAKTY-----RDEHIYPYMY 323
DB 290 FMGVMS-----SCNNPAVNWFHKTIGNFKPYFPWNAPPNEN----- 326
QY 324 LAGYHCRNRNRVREALQAWADTATVIQDYNCREDEEYKEFFEVDVNPILKEA----- 379
DB 327 --PYHAR-RGIKE-----HVIQNAFRKAIQIG 350
QY 380 ASLLEAGEERPGSQSGTQS-----QGSALQ-----DPECFALLRFDYDGICKWEESPT 429

DB 351 LSLMTAGGEPGRGDATSETSSDEDTGQSDVELESDD-----LPYIDP--NMEPVQQR 403
QY 430 PVLHVG-----WATFLVQSLGRFGQVRQKRVIRVREAAEAEEEPWGEAREG 478
DB 404 PYMEFVSVPKPKRKLPPWPTPKTHPVKRTNVKTSR-----SDKAEASQSTPE----- 450
QY 479 RRRGPPRES-----KP-----EPPPPKKP----- 498
DB 451 -RPGSEQSSVTVEPAHTPVPMPVILHQPVPKVPKPTPPPSRRRGACVYVDDDD 509
QY 499 ---ALDKGLGTGQGVASGP-----PRKPPGTAGTARGPE--GGSTA 535
DB 510 VIEVIDVETTESSSVSQPNKPHRKHQDGFQSRGRKRAAPTVPSTGPPAVGPPAA 569
QY 536 QVPA---PAASPPPEGP 549
DB 570 GPPAAGPPAAGPPAAGP 586
RESULT 7
PRP1_HUMAN STANDARD; PRT; 331 AA.
AC P04280;
DT 20-MAR-1987 (Rel. 04, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE SALIVARY PROLINE-RICH PROTEIN PRECURSOR (CLONES CP3, CP4 AND CP5)
DE [CONTAINS: BASIC PEPTIDE IB-6; PEPTIDE P-H].
GN PRB1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85289325; PubMed=2993301;
RA Maeda N., Kim H.-S., Azen E.A., Smithies O.;
RT "Differential RNA splicing and post-translational cleavages in the
RT human salivary proline-rich protein gene system.";
RL J. Biol. Chem. 260:11123-11130(1985).
RN [2]
RP SEQUENCE OF 214-331.
RX MEDLINE=86243355; PubMed=3521730;
RA Kaufman D., Hofmann T., Bennick A., Keller P.;
RT "Basic proline-rich proteins from human parotid saliva: complete
RT covalent structures of proteins IB-1 and IB-6.";
RL Biochemistry 25:2387-2392(1986).
RN [3]
RP SEQUENCE OF 276-331.
RX MEDLINE=84161824; PubMed=6671974;
RA Saitoh E., Isemura S., Sanada K.;
RT "Further fractionation of basic proline-rich peptides from human
RT parotid saliva and complete amino acid sequence of basic proline-rich
RT peptide P-H.";
RL J. Biochem. 94:1991-1997(1983).
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC EMBL; K03204; AAA60185.1; -
DR EMBL; K03205; AAA60186.1; -
DR EMBL; K03206; AAA60187.1; -
DR PIR; A03291; PIHUB6.
DR PIR; C25372; C25372.
DR MIM; 168730; -
KW Repeat; Parotid gland; Multigene family; Saliva; Signal.
FT SIGNAL 1 16

```

FT CHAIN 214 331 PEPTIDE IB-6.
FT CHAIN 276 331 PEPTIDE P-H.
FT VARIAT 106 238 MISSING (IN CLONE CP-4).
FT VARIAT 106 258 /FTIG-VAR_005561.
FT VARIAT 106 276 /FTIG-VAR_005562.
FT CONFLICT 276 A -> S (IN REF. 2 AND 3).
SQ SEQUENCE 331 AA: 32596 MW: 34818F8EBA39751 CRC64;

Query Match 3.9%; Score 123; DB 1; Length 331;
Best Local Similarity 38.4%; Pred. No. 0.35;
Matches 33; Conservative 10; Mismatches 31; Indels 12; Gaps 5;

QY 469 EPMGEAREGRGRPRE--SKPEPPPPKPKALDKGLGTGGQGVAPPPKPPQTVAGTA 526
Db 46 KPGQPPPPKPGQPPGGKNGKPGPPPKP---QG-PPQGDKSRSPSPKPGKPGPP 101
QY 527 RGPGGSTAQVPAP-----AASPPPEG 548
Db 102 --PGGNQPGQPPPPKPGKPGQPPPG 125

RESULT 8
SEPA_EMENI STANDARD; PRT: 1790 AA.
AC P78621;
DT 01-OCT-2000 (Rel. 40, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DT 01-OCT-2000 (Rel. 40, Last annotation update)
DE CYTOKINESIS PROTEIN SEPA (FHL/2 PROTEIN).
GN SEPA.
OS Emericella nidulans (Aspergillus nidulans).
OC Eukaryota; Fungi; Ascomycota; Eurotiales; Trichocomaceae; Emericella.
OX NCBI_TaxID=5072;
RN [1]
RP MEDLINE=97361829; PubMed=9218790;
RA Harris S.D., Hamer L., Sharpless K.E., Hamer J.E.;
RT "The Aspergillus nidulans sepa gene encodes an Fhl/2 protein involved
in cytokinesis and the maintenance of cellular polarity.";
RL EMBO J. 16:3474-3483(1997).
CC -1- FUNCTION: INVOLVED IN CYTOKINESIS.
CC -1- DOMAIN: DFRS ARE REGULATED BY INTRAMOLECULAR GBD-DAD BINDING WHERE
RHO-GTP ACTIVATES THE DFRS BY DISRUPTING THE GBD-DAD INTERACTION
(BY SIMILARITY).
CC -1- SIMILARITY: CONTAINS 1 GTPASE-BINDING DOMAIN (GBD).
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOMOLOGY 1 (FHL) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOMOLOGY 2 (FH2) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 FORMIN HOMOMOLOGY 3 (FH3) DOMAIN.
CC -1- SIMILARITY: CONTAINS 1 DRF AUTOREGULATORY DOMAIN (DAD).
CC -1- SIMILARITY: BELONGS TO THE FORMIN HOMOMOLOGY FAMILY. BNI1
SUBFAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```

```

FT DOMAIN 1608 1611 ARG/LYS-RICH (BASIC).
SQ SEQUENCE 1790 AA: 198628 MW: 0AB355EC1884D05F CRC64;

Query Match 3.8%; Score 122.5; DB 1; Length 1790;
Best Local Similarity 21.4%; Pred. No. 3.3;
Matches 110; Conservative 47; Mismatches 187; Indels 171; Gaps 24;

QY 74 GLTYFPVADLSIIAALYAREFTAQIRGAVDLSLYPREGVSSRELVLKKVSDVIVNLSLSKY 133
Db 680 GLRMYQLVDAML-----SYVAMDRRLPDLDL--RQG-----LTFVQSLL----- 717
QY 134 FKDRHQISLFSFITGTLKSSGVAFVAGCAQLGRDVLHLSADHAWVFGPNG--- 190
Db 718 --DRLHTDAEARRAYDESLEARGIAEA-----ALAEKD-----EMKAQVELCADGLVR 763
QY 191 -----EQTAETVTHCKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMWCAI 243
Db 764 KLOQIEEQTGIIELQSRQNE-----MLKAEIADYOR----- 795
QY 244 NPSTDLHTDSLELLLOOLQOKLLWLLY-----DLGHLERYPMALGNLADLDEL 289
Db 796 -----LRAQELQRLNELETRLYMLRDAQDIAASNAKSNKNGEATDPAHMIRGLDREKL 850
QY 290 EPTPGRPDPLTYHKGITASAKTYRDEHIYPMYLAGYH---CRNRNVREALQAWADTAT 346
Db 851 -----LTRLKQLERTKQFKLEG-----KVMGQHPDRLRELREQMDGDAGP-- 894
QY 347 VIQDYNCREDEEYKEFFEVANDVIPNL-LKEAASLLEAGEERPGESQGTQSQALQ 405
Db 895 -----REAFE--EQARLNLSLNPVGSYVRKTYIQGMEDTATEELGQT-- 935
QY 406 DPCEFAHLLRFYDGICKWEBSGTPVLHVGWATFLVQSLGRFEGOVROKRVISR----- 460
Db 936 DDEVYAKARLVD-----LHRPRMDPEQATGLL-----GEIAAKVPKIDADDAKD 980
QY 461 EAEAAAEAEPEWGEAREGRRRG-----PRRESKPEEPPPPKPKALDKGLGTGGQAVSG 513
Db 981 EGRPTSEQPAEGAATKGDQGVDDTVAVDKATAAP--PPPPPPPAIPIG---SGAAPP 1035
QY 514 PPRKPPCTVAGTARGPEGSGTAQVPAPAAASPPPEG 548
Db 1036 PPPPPPPP-----PPGAGAAPPPIPPPPPPPPG 1063

RESULT 9
PRP2_RAT PRP2_RAT STANDARD; PRT: 172 AA.
AC P10164;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE ACIDIC PROLINE-RICH PROTEIN PRP25 PRECURSOR (FRAGMENT).
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=101116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Parotid gland;
RX MEDLINE=86033799; PubMed=3840480;
RA Clements S., Mehansho H., Carlson D.M.;
RT "Novel multigene families encoding highly repetitive peptide
sequences. Sequence analyses of rat and mouse proline-rich protein
cDNAs."
RT cDNAs.
RL J. Biol. Chem. 260:13471-13477(1985).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL Outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).

```


CC or send an email to license@isb-sib.ch).

CC EMBL; M11901; AAA41948.1; -
DR PIR; B29149; B29149
KW Repeat; Parotid gland; Multigene family; Saliva; Signal.
FT SIGNAL 1 ?
FT CHAIN ? >172 ACIDIC PROLINE-RICH PROTEIN PRP25.
FT NON_TER 172
SQ SEQUENCE 172 AA: 17416 MW; F63BFBFD05459D6EA CRC64;

Query Match 3.7%; Score 119; DB 1; Length 172;

Best Local Similarity 34.0%; Pred. No. 0.28;

Matches 34; Conservative 11; Mismatches 43; Indels 12; Gaps 5;

QY 459 SREAAAEAEPEWGEARGRRRGRRESKPEEPPEP-----KKPALDKGLGTGQGAUSGP 514

DB 76 SFQQKPPQPKFGPPPGGPKQPKPPQPKQPPGPPGQPKRPP-OPKPGPPPPGPP 134

QY 515 PRKPPGTAGTARG--PEGGSTAQVAPAP---ASPPPEGP 549

DB 135 QKPP--QAGKQPPPPGPGQPKPPQPGNQGGPPPPGPP 172

RESULT 10

SYN1_HUMAN

AC P17600; 075825; STANDARD; PRT; 705 AA.

DT 01-AUG-1990 (Rel. 15, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 01-OCT-2000 (Rel. 40, Last annotation update)

DE SYNAPSIN I (BRAIN PROTEIN 4.1).

GN SYN1.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

OX NCBI_TaxID=9606;

[1]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RX MEDLINE=90243651; PubMed=21110562;

RA Suedhof T.C.;

RT "The structure of the human synapsin I gene and protein.";

RL J. Biol. Chem. 265:7849-7852(1990).

[2]

RN SEQUENCE OF 1-125 FROM N.A.

RX MEDLINE=90368667; PubMed=2118519;

RA Sauerwald A.; Hoesche C.; Oschwald R.; Killmann M.W.;

RT "The 5'-flanking region of the synapsin I gene. A G+C-rich, TATA- and

CAAT-less, phylogenetically conserved sequence with cell

type-specific promoter function.";

RL J. Biol. Chem. 265:14932-14937(1990).

[3]

RN SEQUENCE OF 1-258 FROM N.A.

RA Graham D.;

RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: NEURONAL PHOSPHOPROTEIN THAT COATS SYNAPTIC VESICLES,

CC BINDS TO THE CYTOSKELETON, AND IS BELIEVED TO FUNCTION IN THE

CC REGULATION OF NEUROTRANSMITTER RELEASE.

CC -1- SUBCELLULAR LOCATION: SYNAPSE.

CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS; SYNAPSIN IA (SHOWN HERE) AND

CC SYNAPSIN IB; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- PTM: SUBSTRATE OF AT LEAST FOUR DIFFERENT PROTEIN KINASES. IT IS

CC PROBABLE THAT PHOSPHORYLATION PLAYS A ROLE IN THE REGULATION OF

CC SYNAPSIN I IN THE NERVE TERMINAL.

CC -1- SIMILARITY: BELONGS TO THE SYNAPSIN FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to license@isb-sib.ch).

CC EMBL; M58378; AAC41930.1; -
DR EMBL; M58321; AAC41930.1; JOINED.
DR EMBL; M58341; AAC41930.1; JOINED.
DR EMBL; M58351; AAC41930.1; JOINED.
DR EMBL; M58353; AAC41930.1; JOINED.
DR EMBL; M58359; AAC41930.1; JOINED.
DR EMBL; M58371; AAC41930.1; JOINED.
DR EMBL; M58372; AAC41930.1; JOINED.
DR EMBL; M58373; AAC41930.1; JOINED.
DR EMBL; M58374; AAC41930.1; JOINED.
DR EMBL; M58375; AAC41930.1; JOINED.
DR EMBL; M58376; AAC41930.1; JOINED.
DR EMBL; M58377; AAC41930.1; JOINED.
DR EMBL; M58378; AAC41931.1; ALT. SEQ.
DR EMBL; M58321; AAC41931.1; JOINED.
DR EMBL; M58341; AAC41931.1; JOINED.
DR EMBL; M58351; AAC41931.1; JOINED.
DR EMBL; M58353; AAC41931.1; JOINED.
DR EMBL; M58359; AAC41931.1; JOINED.
DR EMBL; M58371; AAC41931.1; JOINED.
DR EMBL; M58372; AAC41931.1; JOINED.
DR EMBL; M58373; AAC41931.1; JOINED.
DR EMBL; M58374; AAC41931.1; JOINED.
DR EMBL; M58375; AAC41931.1; JOINED.
DR EMBL; M58376; AAC41931.1; JOINED.
DR EMBL; M58377; AAC41931.1; JOINED.
DR EMBL; M58378; AAC41931.1; JOINED.
DR EMBL; M55301; AAA60608.1; -
DR EMBL; AL009172; CAA15657.1; -
DR PIR; A35363; A35363.
DR PIR; B35363; B35363.
DR HSP; P17599; LAUX.
DR MIM; 313440; -
DR InterPro; IPR001359; -
DR Pfam; PF02078; Synapsin; 1.
DR PROSITE; PS00415; SYNAPSIN_1; 1.
DR PROSITE; PS00416; SYNAPSIN_2; 1.
KW Synapse; Phosphorylation; Neurone; Repeat; Actin-binding;
KW Alternative splicing.

FT DOMAIN 1 28

FT DOMAIN 29 112

FT DOMAIN 113 420

FT DOMAIN 421 655

FT DOMAIN 656 705

FT MOD_RES 9

FT MOD_RES 568

FT MOD_RES 605

FT VARSPLIC 661

FT VARSPLIC 670

FT VARSPLIC 705

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

FT CONFLICT 138

Query Match 3.7%; Score 118; DB 1; Length 705;

Best Local Similarity 20.8%; Pred. No. 2;

Matches 132; Conservative 67; Mismatches 213; Indels 222; Gaps 34;

QY 66 SPADPPGGLTYFPVADLSI---IAALYARFTAQIRGAVDLSLYPREGGVSSRELV---118

DB 65 AFSPGSSGGGGFSSLSNAVKQTAAATAATFSQVGG---SGGAGRGGAASRVLLVIDE 121

QY 119 -----KKVSDVIWNSLSRSYFKR---AHTQSLFSTFTGKLDSSGVAFAVGA 164

DB 122 PHTDWAKYFKRKHGGIDIKVEQAEFSDLNVAHANGGFS---VDMVELRNGV---KVWRS 177

QY 165 COALGLRDVHLALSDHVVVFGPNGEQTAETVTHGKGNEDRGQTVN---AGVAERSWL 221

DB 178 LKP-----DFVLIHQHAFSM-----ARNGDYRSLVIGIYQAGIPSVNSL 216

QY 222 YLKGSYMRCDRKMEVAFMVCAINPSIDLH-----TDSLELL-----QLOOKLLWL 267

Db 217 H--SVNFCDPKPVFAQMV-----RLHKLGTEEPFLIDQTFYPNHKMLSSSTYPVV 267
QY 268 YDLGHLB--RYPMALGNLADLELETPPGRPDLTLYHKGIASAKTYRDEHIYPNYLA 325
Db 268 VKMGHAHSGMGKVKVDNQHDQDIASV-----VALTKYATAE---PFID-A 310
QY 326 GYHCRNRNREALQA-----WADTATVI----- 348
Db 311 KYDRVQKIQONKAYMRTSVSGNWKNTGCSAMLEQIAMSDRYKLWVDTCSEIFGGLDIC 370
QY 349 -----QDY-----NYCREDEEYKEFEVANDVIPNL----- 375
Db 371 AVEALHGCKDGRHIEVVGSSMPLIGHQDQEDKOLIVEL--VYNKMAQALPRQQRDASP 428
QY 376 -----LKEASNLLEAG---EERPEQ-----SQGTQSQSALQD---PE 408
Db 429 GRGSHGQTPSGALPLGRQTSQAPPAQORPPPGGPPGPGFQROGPPPLQQRPPPPQ 488
QY 409 CFAHLRFYDGICKWEGSPTPVLVHGWATFLVQSLGR-----FEGOVKRVKRVISREAEA 464
Db 489 GQOHL-----SLGP-PAGSLPQORLPSPTSAPQOPASQAAPPTQGGQGRS-RPVAGGPGA 542
QY 465 AEAEPMGEAREGRRGP---RRESKPEPPPPKPKALDKGLGTGCGAVSGVPPRPPGT 521
Db 543 PPAARPPASPQO-ROAGPPQATROTSVSGPAPPKA---SGAPPGGQOQRPQKPPGP 597
QY 522 VAGTARCEGGSTAQVPAPASPP-----PEGP 549
Db 598 AGPTQASQAG-----VPVPTGPTTQQPRPSGP 626
RESULT 11
PRP2_HUMAN
ID PRP2_HUMAN STANDARD; PRT; 251 AA.
AC P02812;
DT 21-JUL-1986 (Rel. 01, Created)
DT 13-AUG-1987 (Rel. 05, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE SALIVARY PROLINE-RICH PROTEIN PRECURSOR (CLONE CP7) [CONTAINS: BASIC
PEPTIDE P-F] (FRAGMENT).
GN PRB2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85289325; PubMed=2993301;
RA Maeda N., Kim H.-S., Azen E.A., Smithies O.;
RT "Differential RNA splicing and post-translational cleavages in the
human salivary proline-rich protein gene system.";
RL J. Biol. Chem. 260:11123-11130(1985).
RN [2]
RP SEQUENCE OF 134-194.
RX MEDLINE=83265674; PubMed=6874669;
RA Saitoh E., Isemura S., Sanada K.;
RT "Complete amino acid sequence of a basic proline-rich peptide, P-F,
from human parotid saliva";
RL J. Biochem. 93:883-888(1983).
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; K03208; AAA60189.1;
DR PIR; A03294; PIHUPF.
DR PIR; E25372; E25372.
DR MIM; 168730;

KW Repeat; Parotid gland; Multigene family; Saliva.
FT NON_TER 1 1
FT CHAIN 134 194 BASIC PEPTIDE P-F.
FT REPEAT <1 9
FT REPEAT 10 71
FT REPEAT 72 133
FT REPEAT 134 195
SQ SEQUENCE 251 AA; 24641 MW; D779F590C0EBF30B CRC64;
Query Match 3.7%; Score 117.5; DB 1; Length 251;
Best Local Similarity 32.3%; Pred. No. 0.57;
Matches 31; Conservative 9; Mismatches 25; Indels 31; Gaps 5;
QY 469 EPMGEAREGRRRGP---SKPEEPPPPKPKALDKGLGTGCGAVSGPP-----R 516
Db 25 QPGPPPPKPKPGPPPGGNGKPGPPPPKPK-----OGPPPGDNKSOSAR 71
QY 517 KPFGTVAGTARGEGGSTAQVPAP---AASPPPEG 548
Db 72 SPFGKPGGP--PQGNQPGPPPPKPKPGPPPG 105
RESULT 12
PRPB_HUMAN
ID PRPB_HUMAN STANDARD; PRT; 1581 AA.
AC Q15648; O75447; O43810;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-OCT-2000 (Rel. 40, Last sequence update)
DE PEROXISOME PROLIFERATOR-ACTIVATED RECEPTOR-ASSOCIATED PROTEIN (PPAR
BINDING PROTEIN) (THYROID HORMONE RECEPTOR INTERACTING PROTEIN 2) (TRIP2)
DE COMPONENT TRAP220) (THYROID RECEPTOR INTERACTING PROTEIN 2) (TRIP2)
DE (P53 REGULATORY PROTEIN RB18A).
GN PPARBP OR TRIP2 OR TRAP220.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 157-168; 943-952 AND 1432-1442.
RX MEDLINE=98318590; PubMed=9633119;
RA Yuan C.X., Ito M., Fondelli J.D., Fu Z.Y., Roeder R.G.;
RT "The TRAP220 component of a thyroid hormone receptor-associated
protein (TRAP) coactivator complex interacts directly with nuclear
receptors in a ligand-dependent fashion";
RL Proc. Natl. Acad. Sci. U.S.A. 95:7939-7944(1998).
RN [2]
RP SEQUENCE OF 16-1581 FROM N.A.
RC TISSUE=Heart;
RX MEDLINE=98105695; PubMed=9444950;
RA Drane P., Barel M., Balbo M., Frade R.;
RT "Identification of RB18A, a 205 kDa new p53 regulatory protein which
shares antigenic and functional properties with p53";
RL Oncogene 15:3013-3024(1997).
RN [3]
RP SEQUENCE OF 622-711 FROM N.A.
RX MEDLINE=95295737; PubMed=7776974;
RA Lee J.W., Choi H.-S., Gyuris J., Brent R., Moore D.D.;
RT "Two classes of proteins dependent on either the presence or absence
of thyroid hormone for interaction with the thyroid hormone
receptor";
RL Mol. Endocrinol. 9:243-254(1995).
CC -!- FUNCTION: INTERACTS WITH THYROID HORMONE RECEPTORS TO REGULATE
CC NUCLEAR RECEPTOR-MEDIATED TRANSCRIPTION. THE PROTEIN HAS BEEN
CC SHOWN TO BIND DNA AND P53 PROTEIN.
CC -!- SUBUNIT: INTERACTS WITH THE LIGAND BINDING DOMAIN OF THE THYROID
CC RECEPTOR (TR). REQUIRES THE PRESENCE OF THYROID HORMONE FOR ITS
CC INTERACTION. HAS THE ABILITY TO SELF-OLIGOMERIZE.
CC -!- ALTERNATIVE PRODUCTS: 2 ISOFORMS; A FORM (SHOWN HERE) AND A SHORT
CC FORM; ARE PRODUCED BY ALTERNATIVE SPLICING. THE SHORT FORM
CC CONTAINS A FRAMESHIFT AFTER GLY-216 AND A STOP CODON AT POSITION
CC 263.

-!- TISSUE SPECIFICITY: EXPRESSED IN ALL TISSUES EXCEPT THE KIDNEYS,
BUT LEVELS ARE HIGHEST IN THE HEART.

-!- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-16 IS THE INITIATOR.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation at the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/) or send an email to license@isb-sib.ch).

```

EMBL; AF055994; AAC39854.1; -.
EMBL; Y13467; AAC73867.1; -.
EMBL; L40366; AAC41736.1; -.
MIN; 604311; -.
DNA-binding; Transcription regulation.
2 X REPEATS OF L-X-X-L-L.
1.
2.
DOMAIN 604 649
REPEAT 604 608
REPEAT 645 649
REPEAT 1099 1111
DOMAIN 1269 1279
POLY-SER.
POLY-SER.
CONFLICT 86 86
CONFLICT 147 147
CONFLICT 471 472
CONFLICT 543 545
CONFLICT 563 563
CONFLICT 573 573
CONFLICT 673 673
CONFLICT 702 708
CONFLICT 721 721
CONFLICT 1388 1388
SEQUENCE 1581 AA; 168517 MW; DD726AFA6895AF57 CRC64
R -> G (IN REF. 2).
F -> S (IN REF. 2).
DS -> GL (IN REF. 2).
LAG -> ASS (IN REF. 2).
P -> S (IN REF. 2).
T -> A (IN REF. 2).
S -> F (IN REF. 3).
MISSING (IN REF. 3).
K -> N (IN REF. 2).
S -> G (IN REF. 2).

```

Query Match 3.7%; Score 117.5; DB 1; Length 1581;
Best Local Similarity 22.4%; Pred. No. 5.9;
Matches 72; Conservative 33; Mismatches 128; Indels 89; Gaps 13;

QY	341	WADTAVIODYNCREDEETIKKEPFEVANDVIPNLLKEAASLLEAGEERPGQSQ-----	395
Db	835	YTPADLIADAAGSPSSDPTNHFFHDGVDENPDLL-NSQSQSGEEYFDESSQGDND	893
QY	396	---GTQSQ-----	416
Db	894	DFKGFASQALNTLGVPMWLGNDGNETKFKGNNOADTVDFSIIISVAGKALAPADLMEH----	949
QY	417	YDGICKWEQSPPTVLHVGWATFLVQSLGRFEGOVQKRYIVSREAAEAEEPPWGEAR	476
Db	950	-----HSGSQGPLTTG-----DLGKERTQKRVK---EGNGTSTNLTSGP-GLDSK	991
QY	477	EGR- RGPRESKEPPPPKPPALDKGLGTGQAVSGPPKPPGTVAGTARGPEGSTA	535
Db	992	PKKRSRTPSNDGSKDKPPKKKADTEGKSPSHSS-SNRPTTP-TSTGSKSPGSAGRS	1041
QY	536	QVPAPASPP-----PEGPVLTPQSEKMKMKELLVATKINSSAIK-----LQL	579
Db	1050	QTPPGVATPPPKITIQIPKTVMW---QKPSHSHQYTSGSGVSSSGSKSHSHSSSSSS	1100
QY	580	TAOSQVMKKQKVSTPESDYTL	601
Db	1107	SASTSGMKSSKSESSSKLS	1128

RESULT 13	MACS_MOUSE	STANDARD;	PRT;	308 AA.
ID	MACS_MOUSE			
AC	P26645;			
DT	01-AUG-1992	(Rel. 23, Created)		
DT	01-AUG-1992	(Rel. 23, Last sequence update)		
DT	01-JUL-1993	(Rel. 26, Last annotation update)		
DE	MYRISTOYLATED ALANINE-RICH C-KINASE SUBSTRATE (MARCKS).			
GN	MACS.			

Mus musculus (Mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
NCBI_TaxID=10990;
(1)
SEQUENCE FROM N.A.
RC TISSUE-Macrophage;
RP MEDLINE=91172836; PubMed=2006186;
RX Seykora J.T., Ravetch J.V., Adreem A.;
RA "Cloning and molecular characterization of the murine macrophage '68-
RT kda' protein kinase C substrate and its regulation by bacterial
RT lipopolysaccharide.";
RN Proc. Natl. Acad. Sci. U.S.A. 88:2505-2509(1991).
RL [2]
SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RC TISSUE-Fibroblast;
RP MEDLINE=91330872; PubMed=1868832;
RX Brooks S.F., Herget T., Erusalimsky J.D., Rozenqurt E.;
RA "Protein Kinase C activation potentially down-regulates the expression
RT of its major substrate, 80K, in Swiss 3T3 cells.";
RN EMBO J. 10:2497-2505(1991).
RL [3]
PARTIAL SEQUENCE.
RC STRAIN-SWISS; TISSUE-Fibroblast;
RP MEDLINE=90346162; PubMed=2384168;
RX Brooks S.F., Erusalimsky J.D., Totty N.F., Rozenqurt E.;
RA "Purification and internal amino acid sequence of the 80 kDa protein
RT kinase C substrate from Swiss 3T3 fibroblasts. Homology with
RN substrates from brain.";
RL FEBS Lett. 268:291-295(1990).
CC -!- FUNCTION: MARKS IS THE MOST PROMINENT CELLULAR SUBSTRATE FOR
CC PROTEIN KINASE C. THIS PROTEIN BINDS CALMODULIN, ACTIN, AND
CC SYNAPSIN. MARKS IS A FILAMENTOUS (F) ACTIN CROSS-LINKING PROTEIN.
CC -!- TISSUE SPECIFICITY: BRAIN, SPLEEN, LESS IN KIDNEY AND HEART, AND
CC VERY LOW LEVELS IN LIVER.
CC -!- INDUCTION: BY LIPOPOLYSACCHARIDE.
CC -!- PTM: PHOSPHORYLATION BY PKC REPLACES MARKS FROM THE MEMBRANE. IT
CC ALSO INHIBITS THE F-ACTIN CROSS-LINKING ACTIVITY.
CC -!- SIMILARITY: BELONGS TO THE MARKS FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC

Query Match 3.7%; Score 117; DB 1; Length 308;
Best Local Similarity 26.3%; Pred. No. 0.79;
Matches 54; Conservative 21; Mismatches 86; Indels 44; Gaps 8;

search completed:
Job time: 199 sec

search completed:
Job time: 199 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:14:32 ; Search time 19.39 Seconds
(without alignments)
2396.414 Million cell updates/sec

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLPLRSIDVV.....KVSTPSDYTLFLKQRKGL 610
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 219241 seqs, 76174552 residues

Total number of hits satisfying chosen parameters: 219241

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR_58:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	126.5	4.0	813	2 T02672	hypothetical prote
2	125.5	3.9	1025	2 S34839	collagen alpha 1(V
3	124.5	3.9	872	2 S33015	hypothetical prote
4	124.5	3.9	992	2 A31666	hypothetical prote
5	124	3.9	392	1 PIH066	salivary proline-r
6	123.5	3.9	1560	2 T02885	peroxisome prolif
7	123.5	3.9	3436	2 S55659	tegumentary protei
8	122.5	3.8	980	2 S54986	regulatory protein
9	119	3.7	172	2 B29149	proline-rich prote
10	119	3.7	705	2 A35363	synapsin I splice
11	118.5	3.7	117	2 D40750	proline-rich prote
12	118	3.7	7962	2 I38346	elastic titin - hu
13	117.5	3.7	251	1 PIHUPF	salivary proline-r
14	117	3.7	309	2 A39169	myristylated alani
15	117	3.7	552	2 T08148	proline-rich myros
16	116.5	3.7	223	2 A42817	proline-rich prote
17	116.5	3.7	2109	2 I38414	transcription fact
18	116	3.6	704	2 A30411	synapsin Ia - rat
19	116	3.6	726	2 S18208	rabphilin-3A-inter
20	116	3.6	2774	2 A43359	microtubule-associ
21	115	3.6	128	2 D38355	basic proline-rich
22	115	3.6	1237	2 T45070	protein kinase hom
23	114.5	3.6	479	2 S10127	muscarinic acetyl
24	114.5	3.6	637	2 T04552	hypothetical prote
25	114	3.6	492	2 A41907	methyl-Cpg-binding
26	113.5	3.6	1838	1 CGH11V	collagen alpha 1(V
27	113	3.5	677	2 T00369	hypothetical prote
28	113	3.5	706	2 E30411	synapsin Ia - bovi
29	113	3.5	1486	1 B40333	collagen alpha 1(I

30	112.5	3.5	310	1 PIHUSD	salivary proline-r
31	112	3.5	188	2 JH0481	basic proline-rich
32	111.5	3.5	2453	2 S60254	nuclear receptor c
33	111	3.5	686	2 A38235	microtubule-associ
34	111	3.5	695	2 T36007	probable transketo
35	111	3.5	744	2 T35192	probable ABC trans
36	111	3.5	1747	1 A45974	collagen alpha 1(X
37	111	3.5	1857	2 S31212	collagen alpha 1(X
38	111	3.5	1888	2 S78476	collagen alpha 1(X
39	110.5	3.5	441	2 T42694	hypothetical prote
40	110	3.5	479	2 S33776	muscarinic acetyl
41	110	3.5	754	2 A56619	female sterile hom
42	110	3.5	858	2 S15762	neurofilament trip
43	110	3.5	1419	2 A41182	collagen alpha 1(I
44	110	3.5	1494	2 T14355	protein-tyrosine-p
45	110	3.5	1844	2 T51890	related to Nup98-N

ALIGNMENTS

RESULT 1
T02672
hypothetical protein R31449_3 - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 02-Sep-2000 #sequence_revision 02-Sep-2000 #text_change 31-Dec-2000
C:Accession: T02672
R:Lamerdin, J.E.; McCready, P.M.; Skowronski, E.; Adamson, A.W.; Burkhardt-Schultz, K.
J.; Danganan, L.; Poundstone, P.; Christensen, M.; Georgescu, A.; Avila, J.; Liu, S.;
P.; Quan, G.; Kronmiller, B.; Arellano, A.; Montgomery, M.; Ow, D.; Nolan, M.
submitted to the EMBL Data Library, June 1998
A:Authors: Trong, S.; Kobayashi, A.; Olsen, A.S.; Carrano, A.V.
A:Description: Sequence analysis of a 3.5 Mb contig in human 19p13.3 containing a ser
A:Reference number: Z14696
A:Accession: T02672
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-813 <L>AM>
A:Cross-references: EMBL:AC005175; NID:g3253116; PID:g3253120
C:Genetics:
A:Map position: 19
A:Introns: 93/2; 125/2; 174/1; 228/2; 266/3; 331/1; 372/1; 474/3; 637/1; 678/3
A:Note: R31449_3

Query Match		4.0%;	Score 126.5;	DB 2;	Length 813;
Best Local Similarity		20.6%;	Pred. No. 0.81;		
Matches		94;	Conservative	74;	Mismatches 172; Indels 117; Gaps 21;
QY	204	EDRRGQTVNAGVAERSWLYL-KGSYMRCDKMEVAFMVCAINPSIDLHTDSLLELLQLOQK 262			
DB	103	EEKELAKERNKRIQEDNRLLEQKVQLRLERREKA-----MREQELEMLOREKE 151			
QY	263	L-----LW-LLYDLGLHLYPYMALGNLADLELEPTGRRDPDLTLYHKGIASAKTYRDEH 317			
DB	152	AEHFKTWEEDQNFHLQQAALR-----SKIRIDGRAPKPIDLLAKYISAEDDLAVEM 204			
QY	318	IYPYMYLAGYHCNRNVREALQAWADATVIQDYNCRDE-----EIVKEFEVANDVI 372			
DB	205	HEPYTFNLGL-----TVADMEDLLEDIQVMELEQGNKADFWRDMTTTIDEI 252			
QY	373	PNLLKEAASILLAGEERPGESQSGQS-----ALQDPECFAHLRLFYDGI-CKWEESG 427			
DB	253	SKLRK-----LEASGKCPGERREGVNASVSSDVQSVFKGKTYNQLOQVIFQIGIECKIRAGG 307			
QY	428	PTFVLHVGNATFLVQSL-----GRFEGQVQRKRVIVSREAFAAAE-----E 469			
DB	308	PN--LDMGYWESLLQQLRAHMAARLRERHQDVLRLKLYKLKE-QGVSEPLFPILKQE 364			
QY	470	PGEEAREGRRRPRRESPEEPKPKALDKGLGTGGQGVSG-PPRKPPTGVAGTARG 528			
DB	365	P-----QSPRSLEPEDAAPTPPGPSSEG-GPAEAVDGAATPTGDDGDGEGEG 413			

Qy 529 PEGSTAQVPAPAAAPPEGVLTFQSEKMKMKE-----LLVATKINSATKIQ 578
Db 414 -----EGEAVLWEEDLIQSLDDYDAGRYSPRLTAHELPLDAHLE 455
Qy 579 LTAQSQ-VQMKKQKVPSPDYLS-----FLKQRKGL 610
Db 456 PDELQRLQLSRQOLQVTGDAESAEDIFFRRAKEGM 492

RESULT 2
S34839
collagen alpha 1(VI) chain precursor - mouse
C:Species: Mus musculus (house mouse)
C:Date: 10-Dec-1993 #sequence_revision 10-Nov-1995 #text_change 13-Aug-1999
C:Accession: S34839; S31405; S32603; S31403; S32003
R:Bonaldi, P.; Piccolo, S.; Marvulli, D.; Volpin, D.; Marigo, V.; Bressan, G.M.
Matrix 13, 223-233, 1993
A:Title: Murine alpha-1(VI) collagen chain. Complete amino acid sequence and identification submitted to the EMBL Data Library, May 1992
A:Reference number: S34839; MUID:93316904
A:Accession: S34839
A:Molecule type: mRNA
A:Residues: 1-1025 <BOW>
A:Cross-references: EMBL:X56405; NID:g50478; PIDN:CAA47032.1; PID:g50479
R:Bonaldi, P.; Piccolo, S.; Marvulli, D.; Volpin, D.; Bressan, G.M.
submitted to the EMBL Data Library, May 1992
A:Reference number: S31403
A:Accession: S31405
A:Molecule type: DNA
A:Residues: 1-31 <BOW>
A:Cross-references: EMBL:X56406; NID:g51055; PIDN:CAA47033.1; PID:g51056
R:Zhang, R.Z.; Pan, T.C.; Timpi, R.; Chu, M.L.
Biochem. J. 291, 787-792, 1993
A:Title: Cloning and sequence analysis of cDNAs encoding the alpha-1, alpha-2 and alpha-3 chains of human collagen alpha 1(VI) chain. Complete amino acid sequence and identification submitted to the EMBL Data Library, March 1988
A:Reference number: S32603; MUID:93256888
A:Accession: S32603
A:Molecule type: mRNA
A:Residues: 442-673, 'TL', 676-708, 'A', 710-942, 944-959, 'R', 961-1025 <ZHA>
A:Cross-references: EMBL:Z18271; NID:g57955; PIDN:CAA79152.1; PID:g57956
C:Genetics:
A:Gene: COL6A1
C:Superfamily: collagen alpha 1(VI) chain; von Willebrand factor type A repeat homology
C:Keywords: cell binding; coiled coil; extracellular matrix; glycoprotein; heterotrimer; F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-1025/Product: collagen alpha 1(VI) chain #status predicted <MAT>
F;34-212/Domain: von Willebrand factor type A repeat homology <WVAL>

Query Match 3.9%; Score 125.5; DB 2; Length 1025;
Best Local Similarity 21.4%; Pred. No. 1.3;
Matches 94; Conservative 36; Mismatches 135; Indels 175; Gaps 23;

Qy 223 LKGYMRCRKM-----EV-----AFMVCA 242
Db 72 LRDRYRCRNLVWNGALHYSDEVEIRGLTRMPSGRDELKASVDKVGKTYTDCA 131
Qy 243 INPSIDLHSTLELL-----LQQLKLLMLLDLGH-LERYPMALGNLAD-LEELEPTGPR 296
Db 132 IKKGL-----ELLIGSHLKNKYLIVVDGHPLEGYKPCGLEDVANEAK----- 179
Qy 297 DPLTYLHKGASATYRDEHIYPM-YLAGYHCRNRNVREALQAWADTATVIQDYNYCR 355
Db 180 -----HLGIKVFVATPDHLEPRLSIATDHYRRNFTAA-----DWGHSR 221
Qy 356 EDEIYKEFEFVANDVPLNLLKEASLEA-----GEE-- 388
Db 222 DAEEVISQITDITVDMIKNNVEQCCSFCEQAARGPPGPRCDPGYEGRGKPLGEGKE 281
Qy 389 -----RCEQS-----QGTQSQ-GS-----ALQDPCF--AHLIRFYDGI----- 420
Db 282 AGDPGRPGDLGVPVQGMKGKSGRKGSRGKRGKIDGDMKGTGYPGL 341
Qy 421 --CKWE-----EGSPTPVLHVGNATLVQSLGRFEGQVRQKRVISREAEAEAEPPWG 472

Db 342 PGCKSGPFGDIQPGPKGDAG-----AFCKMKGKGAGAD-G 379
Qy 473 EEARERRRRGRPRRESKPEEPPEPP--KKPALDKGLGTGGAVSGPPRKPCTVACTA-RGP 529
Db 380 EAGRPGNSGSGDEGDFGPGPGKEGAGDEGNAGPDGA-----PGRGGPGGERGP 431
Qy 530 EGGSTAQVPAPAAAPPPPEGP 549
Db 432 RG--TPGVRGPRGDPGEAGP 449

RESULT 3
S33015
hypothetical protein - human herpesvirus 4
C:Species: human herpesvirus 4, Epstein-Barr virus
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 23-Aug-1997
C:Accession: S33015
R:Farrell, P.J.
submitted to the EMBL Data Library, March 1988
A:Reference number: S32973
A:Accession: S33015
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-872 <FAR>
A:Cross-references: EMBL:V01555
C:Superfamily: human herpesvirus 4 nuclear antigen EBNA-3C

Query Match 3.9%; Score 124.5; DB 2; Length 872;
Best Local Similarity 20.1%; Pred. No. 1.2;
Matches 112; Conservative 53; Mismatches 157; Indels 235; Gaps 25;

Qy 137 RAHIQSLFSFTGTGKLDSSGVAFVAVGACQALGLRDVH-----LALSIEDH 181
Db 1 QSHLQALSNLILDSGLDTQHILCFVMAARQR--LQDIRGGLVAEGGVGHRHLLTSPSQ 58
Qy 182 AWVFGNGEGQTAEV-----TWHGKGNEDRRGQTVNAGV----- 215
Db 59 SW-----PMGYRTATLRLTPVNRVGADSIMLTATFCQNAARTLNTFSATVTPPHIAGP 114
Qy 216 -----AERSWLYLKGSYMRCDKMEVAFVCAINSIDLHTDSLELLOLQOKLLMLLY 268
Db 115 REQERYAREAEVRLRGKWRRYRIYDLIELCG-----SLHHWQNLLOTEENLLDFVR 169
Qy 269 DLGHLERYPMALGNLADLEELEPTGPRDPLTLTHKGASAKTYV---PDEHIYPYMY 323
Db 170 FMGWS-----SCNNPANYWPHKTIQNFKPYYPWNAPI-NEN----- 206
Qy 324 LAGYHCRNRNVREALQAWADTATVIQDYNYCREDEEYKEFEFVANDVITPMLKEA----- 379
Db 207 --PYHAR-RGIKE-----HVIQNAFRKAQIQG 230
Qy 380 ASLLEAGEERPEGSOQGTOS-----OGSALO-----DPECFAHLRLRFYDGIKKEGSPT 429
Db 231 LSMIATGEPGRDGTSETSSDEDGTGSGDVELESSDDE-----LPYIDP--NNEPVOQR 283
Qy 430 PVLHVG-----WATFLVQSLGRFEGQVRQKRVISREAEAEAEAEAEAEAEAEAEAEAE 478
Db 284 PVMFVSRVPAKPKRPLPMPKTPKTHPVKRTNVKTSR-----SDKAEASTPE----- 330
Qy 479 RRGPRRES-----KP-----EPPPPPKP----- 498
Db 331 -REGPSQSSVTVFAHPTVEMPMVILHQPVPVKKPVKPTPPPPRRRRKRGACVYDDD 389
Qy 499 ---ALDKGLGTGQAVSGP-----PRKPPGTVACTARGPE--GGSTA 535
Db 390 VIEVIDVETTEDSSVSQPNKPHRKHQDQFQSRGRRKRAAPPTVSPDGTGPPAVGPAA 449
Qy 536 QVPA---PAASPPPEGP 549
Db 450 GPAAAGPPAAGPPAAGP 466

RESULT 4

A31666

hypothetical protein - human herpesvirus 4

C:Species: human herpesvirus 4, Epstein-Barr virus

C>Date: 20-Jul-1989 #sequence_revision 20-Jul-1989 #text_change 21-Jul-2000

C:Accession: A31666; A49034; B49034; D49034; E49034; F49034

R:Sawada, K.; Yamamoto, M.; Tabata, T.; Smith, M.; Tanaka, A.; Nonoyama, M.

Virolgy 168, 22-30, 1989

A:Title: Expression of EBNA-3 family in fresh B lymphocytes infected with Epstein-Barr virus

A:Reference number: A31666; MUID:89085606

A:Accession: A31666

A:Status: preliminary; not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 1-992 <SAW>

R:Apolloni, A.; Moss, D.; Stumm, R.; Burrows, S.; Suhrbier, A.; Misko, I.; Schmidt, C.; Eur. J. Immunol. 22, 183-189, 1992

A:Title: Sequence variation of cytotoxic T cell epitopes in different isolates of Epstein-Barr virus

A:Reference number: A49034; MUID:92111623

A:Contents: nuclear antigen EBNA-6

A:Accession: A49034

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 281-300 <APO>

A:Cross-references: GB:S79197; NID:9242871; PIDN:AAB20975.1; PID:9242872

A:Experimental source: type A strain B95-8; type A strain IARC/BL74; type A strain IARC/BL74; type A strain IARC/BL74; type A strain IARC/BL74

A:Note: sequence extracted from NCBI backbone (NCBIN:79200, NCBI:79212)

C:Superfamily: human herpesvirus 4 nuclear antigen EBNA-3C

Query Match 3.9%; Score 124.5; DB 2; Length 992;

Best Local Similarity 20.1%; Pred. No. 1.4;

Matches 112; Conservative 53; Mismatches 157; Indels 235; Gaps 25;

QY 137 RAHQSLFETITKLDSSGVAVVGAQALGLRDVH-----LALSEDH 181

DB 121 QSHQLALNLLDGLDTHQILCFVMAQR--LQDIRGLPLVAEGGVGRHWLLTSPSQ 178

QY 182 AWVYFGNGEQTAEV-----TWKGNEDRRGQTNAVY----- 215

DB 179 SW----PMGYRTATRLTPVNRVAGDSIMLTATGCGNAARTLNTSATVTPPHAGP 234

QY 216 -----AERSWLYKSGWRCRMEVAFVCAINPSIDLTFDLSLELQLOKLLWLLY 268

DB 235 REQERYAREAEVRLRGKQRRYRYYDLELGC-----SLHHWQNLQTEENLLDFVR 289

QY 269 DLGHLERYPMALGNLADLELETPGRRPDLTLYHKGIASAKTY-----RDGHYIPYMY 323

DB 290 FGVMS-----SCNNPAVNWFHKTIGNFRPYPNWAPPNEN----- 326

QY 324 LAGYHCNRRNRREALQAWDATVTDYQNYCREDEEYKEFFEVANDVLPNLKEA---- 379

DB 327 --PYHAR-RGIKE-----HVIQNAFRKAQIQG 350

QY 380 ASLLENGERPGSQSGTQS-----QGSALQ-----DPECFAHLRFYDGIKWEBSGPT 429

DB 351 LSLMATGGEPRGDSATSSDSDGTQGSQDVELESSDDE-----LPYIDP--NNEPVPQOR 403

QY 430 PVLHVG-----WATFLVQSLGREGQVRQVRIVSREAAEAEEPEHGEAREG 478

DB 404 PMFVSVRPVPAKPKPLPWPTTPKTHPVKRTNVKTSR-----SDKAEAQSTPE----- 450

QY 479 RRRGPRRES----KP-----BEPPPPKKP----- 498

DB 451 -RGPSEQSSVTVEPAHPTVPVMPVILLHQPPVPVPPVPPVPPVPPVPPVPPVPPVPPV 509

QY 499 ---ALDKGLGTGGAVSGP-----PRKPPGTVAGTARGPE--GGSTA 535

DB 510 VIEVIDVETTEDSSVSQPNKPKRKHQDGFQSRGRKRAAPPTVSPSDTGPVAVGPPAA 569

QY 536 QVPA---PAASPPPEGP 549

DB 570 GPPAAGPAAGPPAAGP 586

RESULT 5

PIHUB6

salivary proline-rich phosphoprotein precursor PRB1 (large allele) [validated] - huma

N:Contains: peptide IB-1; peptide P-E (peptide IB-9); peptide P-F; peptide P-H

C:Species: Homo sapiens (man)

C>Date: 04-Dec-1986 #sequence_revision 12-Apr-1996 #text_change 08-Dec-2000

C:Accession: B40750; C40750; A0750; C25372; S02127; A03293; A90502; A91974;

R:Azen, E.A.; Latreille, P.; Niece, R.L.

Am. J. Hum. Genet. 53, 264-278, 1993

A:Title: PRB1 gene variants coding for length and null polymorphisms among human sali

A:Reference number: A40750; MUID:93304421

A:Accession: B40750

A:Molecule type: DNA

A:Residues: 35-392 <AZE>

A:Cross-references: GB:S62941

A:Experimental source: subject C.J. (large allele)

A:Accession: C40750

A:Molecule type: DNA

A:Residues: 35-127,'R',129-148,'R',150-151,153-187,'K',189-272,'S',274-336,'S',338-39

A:Cross-references: GB:S62929

A:Experimental source: subject M.V.O. (large allele)

A:Accession: A40750

A:Molecule type: DNA

A:Residues: 35-183,245-270,'Q',272-392 <AZ3>

A:Cross-references: GB:S62928

A:Experimental source: subject C.J. (medium allele)

A:Note: authors translated the codon CAA for residue 272 as Arg

R:Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.

J. Biol. Chem. 260, 11123-11130, 1985

A:Title: Differential RNA splicing and post-translational cleavages in the human sali

A:Reference number: A92492; MUID:85289325

A:Accession: C25372

A:Molecule type: mRNA

A:Residues: 1-183,245-392 <MAE>

A:Cross-references: GB:K03204; NID:gl90485; PIDN:AAA60185.1; PID:gl90486

A:Note: alternatively splice forms lacking portions of the repeat region were also fo

R:Lyons, K.M.; Stein, J.H.; Smithies, O.

Genetics 120, 267-278, 1988

A:Title: Length polymorphisms in human proline-rich protein genes generated by intrag

A:Reference number: S02127; MUID:89121440

A:Accession: S02128

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 35-127,250-273,'R',275-277,'R',279-336,'S',338-392 <LYO>

A:Cross-references: EMBL:X07517

A:Accession: S02127

A:Status: translation not shown

A:Molecule type: DNA

A:Residues: 35-183,245-392 <LY2>

A:Cross-references: EMBL:X07516

R:Kaufman, D.; Hofmann, T.; Bennick, A.; Keller, P.

Biochemistry 25, 2387-2392, 1986

A:Title: Basic proline-rich proteins from human parotid saliva: complete covalent str

A:Reference number: A90502; MUID:86243355

A:Accession: A03293

A:Molecule type: protein

A:Residues: 17-38,AP',41-51,92-148,'R',150-152 <KA2>

A:Note: among nine basic proline-rich peptides isolated from the saliva, this peptide

A:Accession: A90502

A:Molecule type: protein

A:Residues: 275-336,'S',338-392 <KAU>

R:Saitoh, E.; Isemura, S.; Sanada, K.

J. Biochem. 94, 1991-1999, 1983

A:Title: Further fractionation of basic proline-rich peptides from human parotid sali

A:Reference number: A91974; MUID:84161824

A:Contents: P-H

A:Accession: A91974

A:Molecule type: protein

A:Residues: 'S',338-392 <SAI>

R:Azen, E.; Lyons, K.M.; McGonigal, T.; Barrett, N.L.; Clements, L.S.; Maeda, N.; Van

Proc. Natl. Acad. Sci. U.S.A. 81, 5561-5565, 1984

A:Reference number: A94005; MUID:84298176

A>Note: the authors translated the codon GGA for residue 151 as Ala, CCT for residue
residue 163 as Pro, GGC for residue 165 as Gln, CCA for residue 166 as Lys, CCA for re
C:Superfamily: proline-rich protein

Query Match 3.7%; Score 119; DB 2; Length 172;
Best Local Similarity 34.0%; Pred. No. 0.34;
Matches 34; Conservative 11; Mismatches 43; Indels 12; Gaps 5;

QY 459 SREAEEAEPEWCEAREGRRRGPRRESKPEEPPPP-----KKPALDKGLGTGGAVSGP 514
 : : : : : : : : : : : : : : : : | | | | | | |
Db 76 SPQKPQPQPKPGPPPQQKGKPPQGKPPGGGPGGQKRP--QPGRKPGQTTPPGGP 134
 :| | | : : | : : : : : : : : : : | | | | |
QY 515 PRKPPGVAGTARG--PEGGSQAQVAPA---ASPPPEGP 549
 :| | | : : | : : : : : : : : : : | | | | |
Db 135 QQKPP--QAKPKQPGPPGGPQPKPPQGNQOGPPPPGPG 172
 :| | | : : | : : : : : : : : : : | | | | |

RESULT 10
A35363
synapsin I splice form a - human
C:Species: Homo sapiens (man)
C>Date: 17-Aug-1990 #sequence_revision 17-Aug-1990 #text_change 01-Dec-2000
C:Accession: A35363; B35363; A35805
R:Suedhof, T.C.
J. Biol. Chem. 265, 7849-7852, 1990
A>Title: The structure of the human synapsin I gene and protein.
A:Reference number: A35363; MUID:90243651
A:Accession: A35363
A:Molecule type: DNA
A:Residues: 1-705 <SUE>
A:CROSS-references: GB:M58371; GB:J05431
A:Accession: B35363
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-659, 'KASPAQAQP' <SUZ>
A:CROSS-references: GB:M58378; GB:J05431
R:Sauerwald, A.; Hoesche, C.; Oschwald, R.; Kilimann, M.W.
J. Biol. Chem. 265, 14932-14937, 1990
A>Title: The 5'-flanking region of the synapsin I gene. A G-C-rich, TAATA- and CAAT-le
A:Reference number: A35805; MUID:90368667
A:Accession: A35805
A:Molecule type: DNA
A:Residues: 1-125 <SAU>
A:CROSS-references: GB:M55301; NID:g338655; PIDN:AAA60608.1; PID:g553654; GB:J05630
C:Genetics:
A:Gene: GDB:SYN1
A:CROSS-references: GDB:l19606; OMIM:313440
A:Map position: Xpl1.23-Xpl1.23
C:Keywords: actin binding; alternative splicing; phosphoprotein

Query Match 3.7%; Score 119; DB 2; Length 705;
Best Local Similarity 20.8%; Pred. No. 2.2;
Matches 132; Conservative 67; Mismatches 213; Indels 222; Gaps 34;

QY 66 SPAEDPPCGLTYPFPVDLSI---IAALYARFTAQIRCAVDLSLYPREGGV:SRRLV---- 118
 : : | : : : : : : : : : : : : : : : | | : : | :
Db 65 APSPGSGGGGFFSSLSNAVKQTAAANATFSEQVG---SGGAGRGGGAASRVLLVIDE 121
 :
QY 119 -----KKYSDDIWNSLSRSFYKDR---AHIQSLFSFITGTKLDSGVAFVAGA 164
 :
Db 122 PHTDWKYFKCKKHGEDIKVEQAEFDNLNVAHANGFS--VDMEVLNGV--KVRS 177
 :
QY 165 QQALGLRDVHLALSIEDHAWVVFGPNGBOTAETVWHGKNEDRRGQTVN-----AGVAERSWL 221
 :
Db 178 LKP-----DPVLIRQHAFSM-----ARNGDYRSLVIGLQYAGIPSVNSL 216
 :
QY 222 YLKGSYRCRKMEVAFMVCAINPSIDLH-----TDLSLEL-----QLQOKLLWLL 267
 :
Db 217 H--SYNFNDKWPWFQAOW-----RLHKLGITGEFPFLDQTFYNHKEMLSTTVPVV 267
 :
QY 268 YDLGHLE--RYPMALGNLADEELPTGPRPDPLTLVHKGIASAKTYTROEHYVPYWLA 325
 :

Db 268 VKMGHAGSGMGKVKVDNQHDQFDIASV-----VALTKTVATAE---PFID-A 310
Qy 326 GYHCNRNRVREALQA-----WADTATVI----- 348
Db 311 KYDVRVOKIGONYKAYMKRTSVSGNKNKTWTSAMLEQIAMSRYKLWVDTCSEIFGGLDIC 370
Qy 349 -----QDY-----NYCREDEETIYKEFFEVDVNPUL----- 375
Db 371 AVEALHGKDRDHLIEVVGSSMLPGHDQEDQLIVEL--VVKMAQALPRQRORDASP 428
Qy 376 -----LKEAASLLRAG-----EERPGEQ-----SQGTOSQGSALQD-----PE 408
Db 429 GRGSHGQTSPSGALPLGRQTSQQPAGPAPQORPPPGQPGPQPGPQPGPQPGPQPGPQ 488
Qy 409 CFAHLRLRYDGIKWEESPTPLVHVGWATFLVQSLGR----FEGQVQKRVIRVSRAEA 464
Db 489 CQQLH-----SLGR-PAGSPLPQRLPSTSAQQPAPASQAAPTQGGGRQS-RPVAGGFGA 542
Qy 465 AEAEPEWGEAREGRRRG-----RRESKPEEPPPKPKALDKGLGTGQAVSGPPPKPPGT 521
Db 543 PPAARPPASPSQ-RQAGPQATQTSVSGPAPKA-----SGAPPQGOORQGPQKPPGP 597
Qy 522 VAGTARPEGGSTAQVPAPASPP-----PEGP 549
Db 598 AGPTRQASQAG-----PVPRTGPPTTQQPRPSGP 626

RESULT 11
D40750
Query Match 3.7%; Score 118.5; DB 2; Length 117;
Best Local Similarity 40.0%; Pred. No. 0.22; Indels 13; Gaps 6;
Matches 34; Conservative 10; Mismatches 28; Mismatches 13; Gaps 6;
Qy 469 EPWGEAREGRRRGPRR--SKPEEPPPKPKALDKGLGTGQAVSGPPPKPPGTVAGTA 526
Db 12 KPQGGPPPPGKPGQPGPQGGNKGKPGPPPPGKP---QG-PPQGGKSRSPRSPGKPGQPP 67
Qy 527 RGPEGGSTAQVPAPASPP---PEGP 549
Db 68 --PQGGNQPQGP---PSPPGKPGQ 87

RESULT 12
I38346
elastic titin - human (fragment)
C:Species: Homo sapiens (man)
C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 21-Jul-2000
C:Accession: I38346
R:Labelit, S.; Koimerer, B.
Science 270, 293-296, 1995
A:Title: Titins: giant proteins in charge of muscle ultrastructure and elasticity.
A:Reference number: A57430; MUID:96026330
A:Accession: I38346
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-7962 <RES>

A:Cross-references: EMBL:X90569; NID:g1017426; PIDN:CAA62189.1; PID:g1017427
C:Genetics:
A:Gene: GDB:TTN
A:Cross-references: GDB:127867; OMIM:188840
A:Map position: 2q31-2q31

Query Match 3.7%; Score 118; DB 2; Length 7962;
Best Local Similarity 25.8%; Pred. No. 63;
Matches 48; Conservative 25; Mismatches 61; Indels 52; Gaps 8;
Qy 420 ICKWEESPTPLVHVGWATFLVQSLGRFEGQV-RQKRVIRVSRAEA---AEAEPEWGEAE 475
Db 6026 VAKKEAPPAKAPVEQ-----KGVVTEEKITIVTQREESPPPAVPEIP-KKKV 6072
Qy 476 REGRRRPRRESKPEEPPPKPKALDKGLGTGQAVSGPPPKPPCTVACTARGPEGGSTA 535
Db 6073 PEERKVPYRKEE--EVPPPKVPKAL-----PKKP-----PEEKVAV 6107
Qy 536 QVPAPASPPPEGVLTFQSEKMKMKELLVATKINSALKLQTAQSQVOMKKQKYSTP 595
Db 6108 PVPVAKKAPPAEVE-----SKTVVEERKREVAEEKLSFAVQVQVETRUEVSAE 6157
Qy 596 SDYTLS 601
Db 6158 EEWYS 6163

RESULT 13
PIHUPF
salivary proline-rich glycoprotein precursor PRB2 [validated] - human (fragment)
N:Alternate names: basic proline-rich peptide IB-8c precursor; proline-rich protein (N:Contains: basic proline-rich peptide IB-4; basic proline-rich peptide P-4
C:Species: Homo sapiens (man)
C:Date: 15-Nov-1984 #sequence_revision 12-Apr-1996 #text_change 08-Dec-2000
C:Accession: E25372; A60827; A03294; B38355; A38355; F38355
R:Maeda, N.; Kim, H.S.; Azen, E.A.; Smithies, O.
J. Biol. Chem. 260, 11123-11130, 1985
A:Title: Differential RNA splicing and post-translational cleavages in the human salivary proline-rich glycoprotein precursor PRB2 [validated]
A:Reference number: A92492; MUID:85289325
A:Accession: E25372
A:Molecule type: mRNA
A:Residues: 1-251 <MAE>
A:Cross-references: GB:K03208; NID:g190509; PIDN:AAA60189.1; PID:g190510
R:Manula, P.W.; Morley, D.J.; Larsen, S.H.; Karn, R.C.
Biochem. Genet. 26, 165-175, 1988
A:Title: Expression of human salivary protein genes.
A:Reference number: A60827; MUID:88240287
A:Accession: A60827
A>Status: not compared with conceptual translation
A:Molecule type: mRNA
A:Residues: 175-251 <MAM>
R:Saitoh, E.; Isemura, S.; Sanada, K.
J. Biochem. 93, 883-888, 1983
A:Title: Complete amino acid sequence of a basic proline-rich p-peptide, p-P, from human parotid saliva: relationships of the
A:Reference number: A03294; MUID:83265674
A:Accession: A03294
A:Molecule type: protein
A:Residues: 134-194 <SAI>
R:Kauffman, D.L.; Bennick, A.; Blum, M.; Keller, P.J.
Biochemistry 30, 3351-3356, 1991
A:Title: Basic proline-rich proteins from human parotid saliva: relationships of the
A:Reference number: A38355; MUID:91190884
A:Accession: B38355
A:Molecule type: protein
A:Residues: 134-194 <KAU>
A:Experimental source: saliva
A:Note: this peptide, which is closely related to that of peptide p-E, contains three
A:Accession: A38355
A:Molecule type: protein
A:Residues: 10-67, R' <RA2>
A:Accession: F38355

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:17:57 ; Search time 21.22 Seconds
(without alignments)
1429.166 Million cell updates/sec

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLFPLRSDDVV.....KVSTPSDYTLFLKQRKGL 610

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 329605 seqs, 49716248 residues

Total number of hits satisfying chosen parameters: 329605

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Nev.*
1: /cgn2_6/ptodata/1/paa/PCT_NEW_COMB.pep.*
2: /cgn2_6/ptodata/1/paa/US06_NEW_COMB.pep.*
3: /cgn2_6/ptodata/1/paa/US07_NEW_COMB.pep.*
4: /cgn2_6/ptodata/1/paa/US08_NEW_COMB.pep.*
5: /cgn2_6/ptodata/1/paa/US09_NEW_COMB.pep.*
6: /cgn2_6/ptodata/1/paa/US60_NEW_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3185	100.0	610	5	US-09-948-941-267
2	3185	100.0	610	5	US-09-380-337-2
3	3185	100.0	639	5	US-09-948-941-387
4	1604	50.4	341	5	US-09-758-466-660
5	818	25.7	160	5	US-09-864-761-38608
6	147	4.6	994	1	PCT-US01-08631-31789
7	127.5	4.0	994	1	PCT-US01-08631-45699
8	123	3.9	331	5	US-09-538-092-845
9	122	3.8	821	5	US-09-884-001-18
10	118.5	3.7	1115	1	PCT-US01-14827-10666
11	118	3.7	705	6	US-60-317-063-2
12	118	3.7	490	5	US-09-538-092-945
13	118	3.7	711	1	PCT-US01-14826-313
14	118	3.7	1273	1	PCT-US01-08631-43847
15	117.5	3.7	251	5	US-09-538-092-840
16	115.5	3.6	140	5	US-09-834-366-13372
17	115.5	3.6	247	5	US-09-538-092-890
18	115	3.6	227	1	PCT-US01-14827-8891
19	115	3.6	520	5	US-09-902-540-10545
20	114.5	3.6	131	5	US-09-834-366-13382
21	114.5	3.6	206	1	PCT-US01-14827-10568
22	111.5	3.5	1203	1	PCT-US01-08656-10232
23	111.5	3.5	1209	1	PCT-US01-08656-10231
24	111	3.5	276	5	US-09-538-092-889
25	110.5	3.5	591	1	PCT-US01-08631-46126
26	110.5	3.5	591	1	PCT-US01-08631-48469
27	110.5	3.5	754	5	US-09-803-110-10319

28	110	3.5	300	1	PCT-US01-08631-50709	Sequence 50709, A
29	110	3.5	362	1	PCT-US01-14827-11292	Sequence 11292, A
30	110	3.5	511	1	PCT-US01-08631-49855	Sequence 49855, A
31	110	3.5	511	6	US-60-311-261-2486	Sequence 2486, Ap
32	110	3.5	807	1	PCT-US01-08631-45532	Sequence 45532, A
33	109.5	3.4	490	6	US-60-317-063-26	Sequence 26, Appl
34	109	3.4	755	1	PCT-US01-24104-57	Sequence 57, Appl
35	109	3.4	755	5	US-09-919-497-57	Sequence 57, Appl
36	109	3.4	1422	1	PCT-US01-08631-32704	Sequence 32704, A
37	108.5	3.4	481	5	US-09-948-941-353	Sequence 353, App
38	108.5	3.4	490	6	US-60-317-063-27	Sequence 27, Appl
39	108.5	3.4	864	5	US-09-538-092-1268	Sequence 1268, Ap
40	108.5	3.4	877	1	PCT-US01-14827-14448	Sequence 14448, A
41	108	3.4	179	5	US-09-758-440-788	Sequence 788, App
42	108	3.4	839	5	US-09-927-796-36	Sequence 36, Appl
43	108	3.4	1136	1	PCT-US01-14827-10458	Sequence 10458, A
44	108	3.4	1136	1	PCT-US01-08631-44700	Sequence 44700, A
45	108	3.4	1136	1	PCT-US01-08631-46239	Sequence 46239, A

ALIGNMENTS

RESULT 1

US-09-948-941-267
; Sequence 267, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948,941
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 267
; LENGTH: 610
; TYPE: PRT
; ORGANISM: Human
US-09-948-941-267

Query Match	100.0%	Score 3185;	DB 5;	Length 610;
Best Local Similarity	100.0%;	Pred. No. 1.9e-185;		
Matches 610;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MGLKAAQKTLFPLRSDDVVRLFAAELGREGPDLVLLSLVLGFEVHFHAVNRVPTNPVE	60	
Db	1	MGLKAAQKTLFPLRSDDVVRLFAAELGREGPDLVLLSLVLGFEVHFHAVNRVPTNPVE	60	
QY	61	LTFQSPADPPGGLTYFPVADLSIIAALYARETAQIRGAVDLSLYPREGVSSRELVLK	120	
Db	61	LTFQSPADPPGGLTYFPVADLSIIAALYARETAQIRGAVDLSLYPREGVSSRELVLK	120	
QY	121	VSDVINLSLSRSYFKDRAHIQSLFSFTTGTCLDSSGVAFVAVGACQALGLRDVHLALSED	180	
Db	121	VSDVINLSLSRSYFKDRAHIQSLFSFTTGTCLDSSGVAFVAVGACQALGLRDVHLALSED	180	
QY	181	HAWVFGPNGEQTAEVYTHWKGKGNEDRRGQTVNAGVAERSWLYLKSGSYMRCDRKMEVAFV	240	
Db	181	HAWVFGPNGEQTAEVYTHWKGKGNEDRRGQTVNAGVAERSWLYLKSGSYMRCDRKMEVAFV	240	
QY	241	CATNPSTDLTDSLELLLOLQOQKLLWLLYDLGLHLETPMALGNLADLEELEPTGCRPDPPLT	300	
Db	241	CATNPSTDLTDSLELLLOLQOQKLLWLLYDLGLHLETPMALGNLADLEELEPTGCRPDPPLT	300	
QY	301	LYHKGTASAKTYRDEHIYPYMYLAGYHCRNRNVREALQAWADTATVIQDYNVCREDEEI	360	
Db	301	LYHKGTASAKTYRDEHIYPYMYLAGYHCRNRNVREALQAWADTATVIQDYNVCREDEEI	360	
QY	361	YKEFFEYANDVIPNLLKEAASLLEANGERPGEQSQGTQSQSALQDPCEFAHLLRFYDGI	420	

Db 361 YKEFEVANDVIPNLLKEAASLLLEAGEERPEQSQGTQSQSALQDPECFALLRFFYDGI 420
Qy 421 CKWEEGSPVPLVHVGWATFLVQSLGREGOVRQKRVIVSREAEAAAEAEPEWGEAREGRR 480
Db 421 CKWEEGSPVPLVHVGWATFLVQSLGREGOVRQKRVIVSREAEAAAEAEPEWGEAREGRR 480
Qy 481 RGPRESKPEPPPPKPPALDKGLGTGQAVSGPPRPPGTVAGTARGPEGGSTAQVPAP 540
Db 481 RGPRESKPEPPPPKPPALDKGLGTGQAVSGPPRPPGTVAGTARGPEGGSTAQVPAP 540
Qy 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSQVQMKKQKVSTPDSYTL 600
Db 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSQVQMKKQKVSTPDSYTL 600
Qy 601 SFLKRQRKGL 610
Db 601 SFLKRQRKGL 610

RESULT 2
US-09-380-337-2
; Sequence 2, Application US/09380337
; GENERAL INFORMATION:
; APPLICANT: Chandrasekharappa, Settara C.
; Manickam, Pachaiappan
; Collins, Francis S.
; Emmert-Buck, Michael R.
; Lubensky, Larisa V.
; Liotta, Lance A.
; Agarwal, Sunita K.
; Spiegel, Allen M.
; TITLE OF INVENTION: MEN1, the Gene Associated with Multiple
; Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses
; Thereof
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; FILING DATE: 09-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/040,269
; FILING DATE: 05-MAR-1997
; APPLICATION NUMBER: WO PCT/US98/04258
; FILING DATE: 04-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Lockyer, Jean M.
; REGISTRATION NUMBER: 44,879
; REFERENCE/DOCKET NUMBER: 015280-315100US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 2:

US-09-380-337-2
Query Match 100.0%; Score 3185; DB 5: Length 610;
Best Local Similarity 100.0%; Pred. No. 1.9e-185;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 MGLKAAQKTLFPLRSIDDVVRLFAAELGREPEDLVLLSLVLGFVEHFLAVNRVLPNTNVE 60
Db 1 MGLKAAQKTLFPLRSIDDVVRLFAAELGREPEDLVLLSLVLGFVEHFLAVNRVLPNTNVE 60
Qy 61 LTQPSAPAPPPGGLTYFPVADLSIIAALYARFTAQIRGAVDSLISLYPREGVSSRELVKK 120
Db 61 LTQPSAPAPPPGGLTYFPVADLSIIAALYARFTAQIRGAVDSLISLYPREGVSSRELVKK 120
Qy 121 VSDVINWSLSRSYFKDRAHIQSLFSFITGKLDSSGVFAFVAVGACQALGLRDVHLALSED 180
Db 121 VSDVINWSLSRSYFKDRAHIQSLFSFITGKLDSSGVFAFVAVGACQALGLRDVHLALSED 180
Qy 181 HAWVFGPNGEQTAEVTHWKGNDREDRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240
Db 181 HAWVFGPNGEQTAEVTHWKGNDREDRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240
Qy 241 CAINPSIDLHTDLSLELLQLOKLLWLLYDLGHLERYPMALGNLADLEELEPTPGRPDPLT 300
Db 241 CAINPSIDLHTDLSLELLQLOKLLWLLYDLGHLERYPMALGNLADLEELEPTPGRPDPLT 300
Qy 301 LYHKGASAKTYRDEHIYPYMLAGYHCRNRNVREALQAWADTATVIQDYNCREDEEI 360
Db 301 LYHKGASAKTYRDEHIYPYMLAGYHCRNRNVREALQAWADTATVIQDYNCREDEEI 360
Qy 361 YKEFEVANDVIPNLLKEAASLLLEAGEERPEQSQGTQSQSALQDPECFALLRFFYDGI 420
Db 361 YKEFEVANDVIPNLLKEAASLLLEAGEERPEQSQGTQSQSALQDPECFALLRFFYDGI 420
Qy 421 CKWEEGSPVPLVHVGWATFLVQSLGREGOVRQKRVIVSREAEAAAEAEPEWGEAREGRR 480
Db 421 CKWEEGSPVPLVHVGWATFLVQSLGREGOVRQKRVIVSREAEAAAEAEPEWGEAREGRR 480
Qy 481 RGPRESKPEPPPPKPPALDKGLGTGQAVSGPPRPPGTVAGTARGPEGGSTAQVPAP 540
Db 481 RGPRESKPEPPPPKPPALDKGLGTGQAVSGPPRPPGTVAGTARGPEGGSTAQVPAP 540
Qy 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSQVQMKKQKVSTPDSYTL 600
Db 541 AASPPPEGPVLTFOSEKMGKMKELLVATKINSSAIKQLTAQSQVQMKKQKVSTPDSYTL 600
Qy 601 SFLKRQRKGL 610
Db 601 SFLKRQRKGL 610

RESULT 3
US-09-948-941-387
; Sequence 387, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948,941
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 387
; LENGTH: 639
; TYPE: PRT
; ORGANISM: Human
US-09-948-941-387

Query Match 100.0%; Score 3185; DB 5; Length 639;
Best Local Similarity 100.0%; Pred. No. 2.1e-185;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLKAAQKTLFPLRSIDDDVVRFAAELGREGPDVLLSLVLGFEVHFVAVNRVPTNVPE 60
DB 30 MGLKAAQKTLFPLRSIDDDVVRFAAELGREGPDVLLSLVLGFEVHFVAVNRVPTNVPE 89
QY 61 LTFQSPADPPGGGLTFYPVADLSIIAALYARFTAQIRGAVDLSLYPRGGVSSRELVRK 120
DB 90 LTFQSPADPPGGGLTFYPVADLSIIAALYARFTAQIRGAVDLSLYPRGGVSSRELVRK 149
QY 121 VSDVWNSLSRSYFKDRAHIQSLFSITGCKLDSSGVAFAVACQALGLRDVHLALSSED 180
DB 150 VSDVWNSLSRSYFKDRAHIQSLFSITGCKLDSSGVAFAVACQALGLRDVHLALSSED 209
QY 181 HAWVFGPNCEQTAETVTHGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFV 240
DB 210 HAWVFGPNCEQTAETVTHGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFV 269
QY 241 CAINPSIDLHTDSLELLQLOQKLLWLLYDLGHLERYPMALGNLADLEELPTPGRPDPLT 300
DB 270 CAINPSIDLHTDSLELLQLOQKLLWLLYDLGHLERYPMALGNLADLEELPTPGRPDPLT 329
QY 301 LYHKGASAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQDNYCREDEEI 360
DB 330 LYHKGASAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQDNYCREDEEI 389
QY 361 YKEFFEVDNDVTPNLLKEAASLLEAGEERPGSQGTQSGSALQDPECFHALLRFYDGI 420
DB 390 YKEFFEVDNDVTPNLLKEAASLLEAGEERPGSQGTQSGSALQDPECFHALLRFYDGI 449
QY 421 CKWEEGSPTPVLHVGNATFLVQSLGRFEGQVQKRVIVSREAEAAEAEPPWEEAREGR 480
DB 450 CKWEEGSPTPVLHVGNATFLVQSLGRFEGQVQKRVIVSREAEAAEAEPPWEEAREGR 509
QY 481 RGPRESKPEEPPPKKPKALDKGLGTGQGVSGPPKPPGTAGTARGEGGSTAQVPAP 540
DB 510 RGPRESKPEEPPPKKPKALDKGLGTGQGVSGPPKPPGTAGTARGEGGSTAQVPAP 569
QY 541 AASPPEGVPLTFQSEKMKMKELLVATKINSATKQLTAQSQVQMKKQVSTPDSYTL 600
DB 570 AASPPEGVPLTFQSEKMKMKELLVATKINSATKQLTAQSQVQMKKQVSTPDSYTL 629
QY 601 SFLKQRKGL 610
DB 630 SFLKQRKGL 639

RESULT 4
US-09-758-466-660
; Sequence 660, Application US/09758466
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PM036
; CURRENT APPLICATION NUMBER: US/09/758,466
; CURRENT FILING DATE: 2001-01-11
; PRIOR APPLICATION NUMBER: 60/179,065
; PRIOR FILING DATE: 2000-01-31
; PRIOR APPLICATION NUMBER: 60/180,628
; PRIOR FILING DATE: 2000-02-04
; NUMBER OF SEQ ID NOS: 814
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 660
; LENGTH: 341
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (115)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-758-466-660

Query Match 50.4%; Score 1604; DB 5; Length 341;
Best Local Similarity 96.3%; Pred. No. 5.9e-90;
Matches 309; Conservative 1; Mismatches 7; Indels 4; Gaps 2;

QY 291 PTPGRP-DPLTLYHKGASAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQ 349
DB 24 PWPARTHELPLP---QHCPSAKTYRDEHIYPYMYLAGYHCRNRNREALQAWADTATVIQ 80
QY 350 DNYCREDEEIIYKEFFEVDNDVTPNLLKEAASLLEAGEERPGSQGTQSGSALQDPE 409
DB 81 DNYCREDEEIIYKEFFEVDNDVTPNLLKEAASLLEAGEERPGSQGTQSGSALQDPE 140
QY 410 FAHLRFYDGI CKWEEGSPTPVLHVGNATFLVQSLGRFEGQVQKRVIVSREAEAAEAE 469
DB 141 FAHLRFYDGI CKWEEGSPTPVLHVGNATFLVQSLGRFEGQVQKRVIVSREAEAAEAE 200
QY 470 PMGEEAREGRRRGRRRESKPEEPPPKKPKALDKGLGTGQGVSGPPKPPGTAGTARGP 529
DB 201 PMGEEAREGRRRGRRRESKPEEPPPKKPKALDKGLGTGQGVSGPPKPPGTAGTARGP 260
QY 530 EGGSTAQVAPAPASPPPEGPVLTFFQSEKMKMKELLVATKINSATKQLTAQSQVQMK 589
DB 261 EGGSTAQVAPAPASPPPEGPVLTFFQSEKMKMKELLVATKINSATKQLTAQSQVQMK 320
QY 590 QKVSTPDSYTL SFLKQRKGL 610
DB 321 QKVSTPDSYTL SFLKQRKGL 341

RESULT 5
US-09-864-761-38608
; Sequence 38608, Application US/09864761
; GENERAL INFORMATION:
; APPLICANT: Penn, Sharron G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: Aomico-X-1
; CURRENT APPLICATION NUMBER: US/09/864,761
; CURRENT FILING DATE: 2001-05-23
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 2000-02-04
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 2000-08-03
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 2000-10-04
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: PCT/US01/00666
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00667
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00664
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00669
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00665
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00668
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00663
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00662
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00661
; PRIOR FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: PCT/US01/00670

[illegible]

	Query Match	3.7%	Score 118;	DB 5;	Length 705;
	Best Local Similarity	20.8%;	Pred. No. 16;		
	Matches 132;	Conservative	67;	Mismatches 213;	Indels 222; Gaps
Qy	66	SPADPPGCLTYFPVADLSI---	TAALYAREFTAQIRCAVDLSL	YPREGGVSSHELV----	118
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Db	65	APSPGGSGGGFFSLSNVAKQT	TAAAAATFSEQVGG---SGCAGRGCAASRVLLVIDE	1211	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Qy	119	-----KKYSDDIWNLSLSRSYKDR---	AHQSLFSFTCTGLDSSGVAFAVVGA	164	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Db	122	PHTDWAKYFKGKKIHGGIDIKVEQAEE	SDNLVAHANGFES--VDMEVLRNGV--KVVR5	177	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Qy	165	COALGLRDVHLALSIEDHAWVVFGPNGETA	EVTWHGKGNEDRGQTVN---AGVAERSWL	221	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Db	178	LKP -----DRVLITROHAFSM-----	-ARNGDYRSLVIGIQYAGIPSVNSL	216	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Qy	222	YLKGSYMRCDRKMEVAFMVCAINPSIDLH---	TDSLLEL-----OLQOKLWLWL	267	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Db	217	H--SYVNCDRPWTFAQWV-----	RLHHKLGTETPEPLDQTFYPMUKEMLSSTHYPVV	267	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Qy	268	YDLGHLE--RYPMALGNLADLEEPTPGRPDP	LTLYHKGTASAKTYRYRDEHIYPMYLA	325	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Db	268	VKMGHASHGCMKVKNQHDFQDIATSV-----	VALTKTYATAE--PFID-A	310	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:
Qy	326	GHCRCNRNVREALQA-----	WADTATVI-----	348	
	:	:	:	:	:
	:	:	:	:	:
	:	:	:	:	:

Db 311 KYDVRVOKIGONYKAYMRTSVSGNWKNTGTSAMLEQIAMSRYKLWVDRCTSEIFGGGLDIC 370
QY 349 -----QDY-----NYCREDEIYKEFFEYANDVIPNL----- 375
Db 371 AVEALHGKGRDHIIEVVGSSMPLIGHQDQDLVEL--VVNKAQALPRQRDASP 428
QY 376 -----IKEAASLEAG---EERPEQ-----SQGTQSGSALQD---PE 408
Db 429 GRGSHGQTSPGALPLGRQTSQAPAGPAQORPPQGGPQPGPQROGPPLOQRPPQ 488
QY 409 CFAHLRLFDVGICKWEGSPTVLHVGNATFLVQSIGR-----FEGQVROKVRIVSREAEA 464
Db 489 GOQHL-----SGLGP-PAGSPLFLQRLSPSTSAPQQPASQAAPTQGGQGRS-REVAGGPGA 542
QY 465 AEAEPEWGEAREGRRGP---RRESKPEEPPPPKPPALDKGLGTGGGAVSGPPEKPPGT 521
Db 543 PPAARPPASPSPO-ROAGPPQATRTQTSVSGPAPPKA-----SGAPPGGQORQPPQRP 597
QY 522 VAGTARGPEGGSTAQPAPAAASP-----PEGP 549
Db 598 AGPTRQASQAG-----PVPRTGPTTQOPRPSGP 626

RESULT 13
PCT-US01-14826-313
; Sequence 313, Application PC/TUS0114826
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-103
; CURRENT APPLICATION NUMBER: PCT/US01/14826
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; PRIOR APPLICATION NUMBER: 09/577,298<151> 2000-09-22
; PRIOR APPLICATION NUMBER: 09/695,781<151> 2000-10-24
; PRIOR APPLICATION NUMBER: 09/715,869<151> 2000-11-17
; PRIOR APPLICATION NUMBER: 09/775,330<151> 2001-02-01
; NUMBER OF SEQ ID NOS: 864
; SOFTWARE: Custom
; SEQ ID NO 313
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US01-14826-313

Query Match 3.7%; Score 118; DB 1; Length 711;
Best Local Similarity 25.1%; Pred. No. 17;
Matches 52; Conservative 15; Mismatches 62; Indels 78; Gaps 9;
QY 411 AHLLR-FYDGICKWEGSPTVLHVGNATFLVQS-----LGRFEGQVROKVRIVSRE 461
Db 60 SHWARGFQNTCGPRGSPQPPRRPWASRVLQEATNWRAGPLAEVRAEQRKAASQE 119
QY 462 AEAAEAEPEWGEAREGRRGPRRESKPE-----EPPP 494
Db 120 REAKETERK--RRKAGGARRSPGPRPEPRNAPRVAQLAGLPAPLPERLAPVGRAPR 177
QY 495 PKPALDKG---LGTGGGAVSGGPR-----KPPGTVA----- 523
Db 178 SAQPSQSDPGSAWAGPWGRRGPRGPPSYEAHLLLRGSAGTAPRRWRDRPPPVAPPSEYEGPH 237
QY 524 ---GTARGPEGGSTAQP---APAAASP 544
Db 238 RTLGTRGP---GNSQVPTSSAPAAATP 261

RESULT 14
PCT-US01-08631-43847
; Sequence 43847, Application PC/TUS0108631
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc

; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-049
; CURRENT APPLICATION NUMBER: PCT/US01/08631
; CURRENT FILING DATE: 2001-03-30
; PRIOR APPLICATION NUMBER: 09/540,217
; PRIOR FILING DATE: 2000-03-31
; PRIOR APPLICATION NUMBER: 09/649,167
; PRIOR FILING DATE: 2000-08-23
; NUMBER OF SEQ ID NOS: 60736
; SOFTWARE: Custom
; SEQ ID NO 43847
; LENGTH: 1273
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: DOMAIN
; LOCATION: (751)..(765)
; OTHER INFORMATION: Phage integrase family domain identified by eMATRIX,
; OTHER INFORMATION: accession number PF00589B, p-value=1.621e-11, raw score of 16.
; NAME/KEY: misc.feature
; LOCATION: (1)..(1273)
; OTHER INFORMATION: Xaa = X or * as defined in Table 2
PCT-US01-08631-43847

Query Match 3.7%; Score 118; DB 1; Length 1273;
Best Local Similarity 25.1%; Pred. No. 35;
Matches 52; Conservative 15; Mismatches 62; Indels 74; Gaps 9;
QY 411 AHLLR-FYDGICKWEGSPTVLHVGNATFLVQS-----LGRFEGQVROKVRIVSRE 461
Db 60 SHWARGFQNTCGPRGSPQPPRRPWASRVLQEATNWRAGPLAEVRAEQRKAASQE 119
QY 462 AEAAEAEPEWGEAREGRRGPRRESKPE-----EPPP 494
Db 120 REAKETERK--RRKAGGARRSPGPRPEPRNAPRVAQLAGLPAPLPERLAPVGRAPR 177
QY 495 PKPALDKG---LGTGGGAVSGGPR-----KPPGTVA----- 523
Db 178 SAQPSQSDPGSAWAGPWGRRGPRGPPSYEAHLLLRGSAGTAPRRWRDRPPPVAPPSEYEGPH 237
QY 524 ---GTARGPEGGSTAQP---APAAASP 544
Db 238 RTLGTRGP---GNSQVPTSSAPAAATP 261

RESULT 15
US-09-538-092-840
; Sequence 840, Application US/09538092
; GENERAL INFORMATION:
; APPLICANT: Glot, Loic
; APPLICANT: Mansfield, Traci A.
; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
; FILE REFERENCE: 15966-542
; CURRENT APPLICATION NUMBER: US/09/538,092
; CURRENT FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 60/127,352
; PRIOR FILING DATE: 1999-04-01
; PRIOR APPLICATION NUMBER: 60/178,965
; PRIOR FILING DATE: 2000-02-01
; NUMBER OF SEQ ID NOS: 1387
; SOFTWARE: CurapatSeqformatter Version 0.9
; SEQ ID NO 840
; LENGTH: 251
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc.feature
; LOCATION: (0)...(0)
; OTHER INFORMATION: Polypeptide Accession Number P02812
US-09-538-092-840

Query Match 3.7%; Score 117.5; DB 5; Length 251;
Best Local Similarity 32.3%; Pred. No. 4.8;
Matches 31; Conservative 9; Mismatches 25; Indels 31; Gaps 5;
Qy 469 EPWGEAREGRRRGPRRE--SKPEPPPPKPKALDKGLGTGGAVSGPP-----R 516
Db 25 PQGGPPPPGKPPGPPGPGGKPKGPPPPGPP-----QGPPPGDNKKSQAR 71
Qy 517 KPPGTVAGTARGEGGTAQVPAP----AASPPPEG 548
Db 72 SPPGKPPGPP--PQGGNPPGPPPPPPGKPPGPPG 105

Search completed: October 12, 2001, 18:21:16
Job time: 199 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:17:22 ; Search time 90.47 Seconds
(without alignments)
1495.821 Million cell updates/sec

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLFLRLSDVV.....KVSTPSDYTLFLKQRKGL 610

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2803329 seqs, 221847457 residues

Total number of hits satisfying chosen parameters: 2803329

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_AA_Main:*

1:	/cgn2_6/ptodata/2/paa/PCTUS_COMB.pep.*
2:	/cgn2_6/ptodata/2/paa/US06_COMB.pep.*
3:	/cgn2_6/ptodata/2/paa/US07_COMB.pep.*
4:	/cgn2_6/ptodata/2/paa/US080_COMB.pep.*
5:	/cgn2_6/ptodata/2/paa/US081_COMB.pep.*
6:	/cgn2_6/ptodata/2/paa/US082_COMB.pep.*
7:	/cgn2_6/ptodata/2/paa/US083_COMB.pep.*
8:	/cgn2_6/ptodata/2/paa/US084_COMB.pep.*
9:	/cgn2_6/ptodata/2/paa/US085_COMB.pep.*
10:	/cgn2_6/ptodata/2/paa/US086_COMB.pep.*
11:	/cgn2_6/ptodata/2/paa/US087_COMB.pep.*
12:	/cgn2_6/ptodata/2/paa/US088_COMB.pep.*
13:	/cgn2_6/ptodata/2/paa/US089_COMB.pep.*
14:	/cgn2_6/ptodata/2/paa/US090_COMB.pep.*
15:	/cgn2_6/ptodata/2/paa/US091_COMB.pep.*
16:	/cgn2_6/ptodata/2/paa/US092_COMB.pep.*
17:	/cgn2_6/ptodata/2/paa/US093_COMB.pep.*
18:	/cgn2_6/ptodata/2/paa/US094_COMB.pep.*
19:	/cgn2_6/ptodata/2/paa/US095_COMB.pep.*
20:	/cgn2_6/ptodata/2/paa/US096_COMB.pep.*
21:	/cgn2_6/ptodata/2/paa/US097_COMB.pep.*
22:	/cgn2_6/ptodata/2/paa/US098_COMB.pep.*
23:	/cgn2_6/ptodata/2/paa/US060_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3185	100.0	610	12	US-08-865-337-6
2	2778	87.2	564	12	US-08-865-337-1
3	1124.5	35.3	813	23	US-60-171-625-525
4	1124.5	35.3	813	23	US-60-173-464-25061
5	1124.5	35.3	813	23	US-60-191-637-32161
6	1124.5	35.3	813	23	US-60-191-681-25486
7	1124.5	35.3	813	23	US-60-219-005-306
8	818	25.7	160	1	PCT-US01-00863-31783
9	619	19.4	120	23	US-60-196-710-5355
10	462.5	14.5	92	23	US-60-177-646-2142

11	342.5	10.8	72	23	US-60-177-646-2905	Sequence 2905, Ap
12	140.5	4.4	518	23	US-60-167-217-9939	Sequence 9939, Ap
13	140.5	4.4	518	23	US-60-173-464-8029	Sequence 8029, Ap
14	131	4.1	1151	14	US-09-023-905-4	Sequence 4, Appli
15	129	4.1	566	23	US-60-191-637-9944	Sequence 9944, Ap
16	129	4.1	566	23	US-60-191-681-7769	Sequence 7769, Ap
17	127	4.0	834	1	PCT-US01-08117-34	Sequence 34, Appli
18	127	4.0	834	23	US-60-188-986-34	Sequence 34, Appli
19	125.5	3.9	603	23	US-60-230-435-1481	Sequence 1481, Ap
20	123.5	3.9	720	1	PCT-US01-01567-101	Sequence 101, App
21	123.5	3.9	766	1	PCT-US01-01567-100	Sequence 100, App
22	123.5	3.9	973	1	PCT-US01-01567-61	Sequence 61, Appli
23	123.5	3.9	973	1	PCT-US01-11988-1424	Sequence 1424, Ap
24	123.5	3.9	973	22	US-09-833-245-1424	Sequence 1424, Ap
25	121	3.8	456	16	US-09-252-991A-21576	Sequence 21576, A
26	121	3.8	930	23	US-60-280-260-1	Sequence 1, Appli
27	119	3.7	705	18	US-09-488-725A-2473	Sequence 2473, Ap
28	119	3.7	1291	23	US-60-242-679-769	Sequence 769, App
29	118.5	3.7	501	18	US-09-417-507-23911	Sequence 23911, A
30	118.5	3.7	714	18	US-09-488-725A-6045	Sequence 6045, Ap
31	118	3.7	184	16	US-09-270-767-41823	Sequence 41823, A
32	118	3.7	1336	1	PCT-US98-24950-20	Sequence 20, Appli
33	118	3.7	1336	13	US-08-975-424-20	Sequence 20, Appli
34	118	3.7	7962	17	US-09-312-216B-36	Sequence 36, Appli
35	117.5	3.7	124	1	PCT-US97-15799-2	Sequence 2, Appli
36	117.5	3.7	1581	15	US-09-110-517-2	Sequence 3, Appli
37	117.5	3.7	1581	15	US-09-110-517-3	Sequence 2, Appli
38	117.5	3.7	1581	17	US-09-346-231-2	Sequence 2, Appli
39	117.5	3.7	1581	17	US-09-346-231-3	Sequence 3, Appli
40	117.5	3.7	1581	17	US-09-346-231-2	Sequence 2, Appli
41	117	3.7	114	16	US-09-270-767-57068	Sequence 57068, A
42	117	3.7	114	16	US-09-270-849B-181000	Sequence 181000,
43	117	3.7	309	8	US-08-405-175-7	Sequence 7, Appli
44	117	3.7	696	15	US-09-129-668-4	Sequence 4, Appli
45	117	3.7	1442	23	US-60-229-518-325	Sequence 325, App

ALIGNMENTS

RESULT 1
US-08-865-337-6
; Sequence 6, Application US/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 610 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GenBank
CLONE: 1945387
US-08-865-337-6

Query Match 100.0%; Score 3185; DB 12; Length 610;
Best Local Similarity 100.0%; Pred. No. 2.8e-259;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MGLKAAQKTLFPLRSIDDVVRLFAAELGREPDVLVLSVLGFEHFLAVNRVPTNVPE 60
Db 1 MGLKAAQKTLFPLRSIDDVVRLFAAELGREPDVLVLSVLGFEHFLAVNRVPTNVPE 60
QY 61 LTFQSPAPDPGGTLTFPVADLSIIAALYARFTAQIRGAVDLSLYPREGVSSRELKVK 120
Db 61 LTFQSPAPDPGGTLTFPVADLSIIAALYARFTAQIRGAVDLSLYPREGVSSRELKVK 120
QY 121 VSDVINSLRSYFKDRAHIQSLFSFTGTGKLDSSGVAFVAVGACQALGRDVLHLS 180
Db 121 VSDVINSLRSYFKDRAHIQSLFSFTGTGKLDSSGVAFVAVGACQALGRDVLHLS 180
QY 181 HAWVVFPGNGEQTAEVTHWKGNDREGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFV 240
Db 181 HAWVVFPGNGEQTAEVTHWKGNDREGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFV 240
QY 241 CAINPSIDLHTDSLLELLOKLLWLLYDLGLHLERYPMALGNLADLELEPTPGRPDPLT 300
Db 241 CAINPSIDLHTDSLLELLOKLLWLLYDLGLHLERYPMALGNLADLELEPTPGRPDPLT 300
QY 301 LYHKGIAKTYRDEHIYPYMLAGYHCRNRNVRREALQAWADTATVIQDNYNCREDEE 360
Db 301 LYHKGIAKTYRDEHIYPYMLAGYHCRNRNVRREALQAWADTATVIQDNYNCREDEE 360
QY 361 YKEFEVANDVIPNLLKEASLLEAGEERPEQSGQTSQSGSALQDPECFHLLRFYDGI 420
Db 361 YKEFEVANDVIPNLLKEASLLEAGEERPEQSGQTSQSGSALQDPECFHLLRFYDGI 420
QY 421 CKWEEGSPTPVLHVGWATFLVQSLGREGQVRQKRVIVSREAAEAEEPEWGEAREGR 480
Db 421 CKWEEGSPTPVLHVGWATFLVQSLGREGQVRQKRVIVSREAAEAEEPEWGEAREGR 480
QY 481 RGPRESKPEPPPPKPPALDKGLTGCGAVSGPPRPGTGTAGTARGPEGGTAQVPAP 540
Db 481 RGPRESKPEPPPPKPPALDKGLTGCGAVSGPPRPGTGTAGTARGPEGGTAQVPAP 540
QY 541 AASPPPEPVLTFQSEKMGKMKELLVATKINSSAIKQLTAQSOVMKKOKVSTPSPDYTL 600
Db 541 AASPPPEPVLTFQSEKMGKMKELLVATKINSSAIKQLTAQSOVMKKOKVSTPSPDYTL 600
QY 601 SFLKRQRKGL 610
Db 601 SFLKRQRKGL 610
```

RESULT 2

US-08-865-337-1
Sequence 1, Application US/08865337
GENERAL INFORMATION:
APPLICANT: Au-Young, Janice
APPLICANT: Covitz, Peter
APPLICANT: Tang, Y. Tom

APPLICANT: Murry, Lynn E.
TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,337
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 564 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BRAINOT14
CLONE: Consensus
US-08-865-337-1

Query Match 87.2%; Score 2778; DB 12; Length 564;
Best Local Similarity 89.5%; Pred. No. 5.5e-225;
Matches 547; Conservative 2; Mismatches 14; Indels 48; Gaps 4;

```
QY 1 MGLKAAQKTLFPLRSIDDVVRLFAAELGREPDVLVLSVLGFEHFLAVNRVPTNVPE 60
Db 1 MGLKAAQKTLFPLRSIDDVVRLFAAELGREPDVLVLSVLGFEHFLAVNRVPTNVPE 60
QY 61 LTFQSPAPDPGGTLTFPVADLSIIAALYARFTAQIRGAVDLSLYPREGVSSRELKVK 120
Db 61 LTFQSPAPDPGGTLTFPVADLSIIAALYARFTAQIRGAVDLSLYPREGVSSRELKVK 120
QY 121 VSDVINSLRSYFKDRAHIQSLFSFTGTGKLDSSGVAFVAVGACQALGRDVLHLS 180
Db 121 VSDVINSLRSYFKDRAHIQSLFSFTGTGKLDSSGVAFVAVGACQALGRDVLHLS 180
QY 181 HAWVVFPGNGEQTAEVTHWKGNDREGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFV 240
Db 181 HAWVVFPGNGEQTAEVTHWKGNDREGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFV 240
QY 241 CAINPSIDLHTDSLLELLOKLLWLLYDLGLHLERYPMALGNLADLELEPTPGRPDPLT 300
Db 241 CAINPSIDLHTDSLLELLOKLLWLLYDLGLHLERYPMALGNLADLELEPTPGRPDPLT 300
QY 301 LYHKGIAKTYRDEHIYPYMLAGYHCRNRNVRREALQAWADTATVIQDNYNCREDEE 359
Db 301 LYHKGIAKTYRDEHIYPYMLAGYHCRNRNVRREALQAWADTATVIQDNYNCREDEE 359
QY 361 YKEFEVANDVIPNLLKEASLLEAGEERPEQSGQTSQSGSALQDPECFHLLRFYDGI 419
Db 361 YKEFEVANDVIPNLLKEASLLEAGEERPEQSGQTSQSGSALQDPECFHLLRFYDGI 419
```



```
Db 335 IYKEFFEYVANDVIPNLLKEAASLLEAGEERPGESQSQSALQDPECFAHLLRFYDG 394
QY 420 ICKWEGSPPTVLHVGWATFLVQSLGRFEGOVROKVIYVSREAAEAEEPEWCEAREGR 479
Db 395 ICKWEGSPPTVLHVGWAPFLKSLARFGVQVXOKVRIYVDAEAAE-----441
QY 480 RRGPRRESKPEPPPPKPKALDKGLGTGOGAVSGPPRKPPTGTVAGTARGEGGSTAQVPA 539
Db 442 -----KPEPPPPKPKALDKGLGTGOGAVSGPPRKPPTGTVAGTARGEGGSTAQVPA 493
QY 540 PAASPPPEGVLTFSQSEKMKMKELLVATKINSKAIKQLTAQSQVOMKKQKVPSTPSDYT 599
Db 494 PAASPPPEGVLTFSQSEKMKMKELLVATKINSKAIKQLTAQSQVOMKKQKVPSTPSDYT 553
QY 600 LSFLKRQRKGL 610
Db 554 LSFLKRQRKGL 564

RESULT 3
US-60-171-625-525
; Sequence 525, Application US/60171625
; GENERAL INFORMATION:
; APPLICANT: Skupski, Marian
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS THAT SHOW
; TITLE OF INVENTION: HIGH HOMOLOGY TO KNOWN HUMAN DISEASE PROTEINS, AND USES
; FILE REFERENCE: CL000179
; CURRENT APPLICATION NUMBER: US/60/171,625
; CURRENT FILING DATE: 1999-12-23
; NUMBER OF SEQ ID NOS: 579
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 525
; LENGTH: 813
; TYPE: PRT
; ORGANISM: DROSOPHILA
US-60-171-625-525

Query Match 35.3%; Score 1124.5; DB 23; Length 813;
Best Local Similarity 35.4%; Pred. No. 3.4e-85;
Matches 274; Conservative 88; Mismatches 198; Indels 213; Gaps 16;

QY 9 TLPLRSIDDDVRLFAEL-GREEPDVLVLSVLGVFVEHFLAVNRVPTNPVELTFOQSP 67
Db 10 SLFPLKSTADVINLFRALTSIEPDLTLLSIVGVYIELSLTTGEAAQ-----AAQAAA 64

QY 68 APDPGGLTY-----FPVADLSIIAALYARFTAQIRGAVDLSLYPREGV 112
Db 65 AAVAAGDISQATTGGNDIIMGNSVPPVVTHELIALGKFKQT-ILSVVE---KPKPHRQ 120

QY 113 SSRELKVKVSDVWNSLSRSYFKDRAHIQSLFSFITGKLDSSGVAFVAVGACQALGLRD 172
Db 121 ATREVTKVSDVWNSLSRSYFKDRAHLQNLYSYLSGNKLDLFCGVALAVVAGCOLLYGKD 180

QY 173 VHLAISEDHAWVYFGPNEGQTAETVTHGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDR 232
Db 181 VRLAISEDHAWVYFGQKRVETIEVTHGKGSDEKRGQDIRPGIESGSWLYLGLGLAVVCCR 240

QY 233 KMEVAFMVCAINPSIDLHTSDLSLELLQOQKLLWLLDGLHLERYPMALGNLADLEELEPT 292
Db 241 GMEVAACAALNLSLTSNSDCVEAELOQQLLWLLDGLHLKRYPMALGTLGEELEHRT 300

QY 293 Y-CREDEEYKFEFFVANDVIPNLLKEAASLLEAGEERPGESQSQSQSQ-----SALQDP 407
Db 361 YQCRDEEYKFEFFVANDVIPNLLKEAASLLEAGEERPGESQSQSQSQ-----SALQDP 401

QY 408 ECFHLLRFYDGICKWEGSPPTVLHVGWATFLVQSLGRFEGOVROKVIYVSREAAEAEEPEWCEAREGR 457
Db 408 ECFHLLRFYDGICKWEGSPPTVLHVGWATFLVQSLGRFEGOVROKVIYVSREAAEAEEPEWCEAREGR 457
```

```
Db 402 EVFANLLRFYDGICQWEDSLTPIHLGWAKPLVNNITKFDYDIRSQVVIKLPEDLEAEQ 461
QY 458 -----VSREA-EAAEAEPWCEAREGR-----GPRRESKPEEP---493
Db 462 AKAEARAEQAEKAEKAEAGSEAWEGNNNRWATKEERNVLAIGKELSSKPRNSQSSL 521
QY 494 -----PPKKPALDKGL-----504
Db 522 LLINISVLTETTSYLLAVTLFTMLVHGLVSFVAESAMVLYKSKNSLPPTTLADLTAACG 581
QY 505 -----GTGQ-----GAVSGPP-----RKPPGTV 522
Db 582 EKITLPDFLLQGGQPFADQKQPSGGESDPNPELHNNNNNSNNNNNNHNNADKKEAAAT 641
QY 523 AGTARGEGGSTAQVPAAPASPPPEG-----548
Db 642 TTNATTTSGSGTSVOLPVSSEANNAGAQSQVQINDQLGKPOHKEAKKETSDDYDPFE 701
QY 549 -----PVLTFQSEKMKMKELLVATKINSKAIKQLTAQSQVOMKKQKVPSTPS 596
Db 702 IMLKRPVITLYSQMKGLDKLLAELNTHAISLQVTAQSVASRKVRGVEKHS 754

RESULT 4
US-60-173-464-25061
; Sequence 25061, Application US/60173464
; GENERAL INFORMATION:
; APPLICANT: Li, Peter W.D.
; TITLE OF INVENTION: ISOLATED G-PROTEIN COUPLED RECEPTORS,
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING GPCR PROTEINS AND USES
; FILE REFERENCE: CL000173
; CURRENT APPLICATION NUMBER: US/60/173,464
; CURRENT FILING DATE: 1999-12-29
; NUMBER OF SEQ ID NOS: 30269
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 25061
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Drosophila
US-60-173-464-25061

Query Match 35.3%; Score 1124.5; DB 23; Length 813;
Best Local Similarity 35.4%; Pred. No. 3.4e-85;
Matches 274; Conservative 88; Mismatches 198; Indels 213; Gaps 16;

QY 9 TLPLRSIDDDVRLFAEL-GREEPDVLVLSVLGVFVEHFLAVNRVPTNPVELTFOQSP 67
Db 10 SLFPLKSTADVINLFRALTSIEPDLTLLSIVGVYIELSLTTGEAAQ-----AAQAAA 64

QY 68 APDPGGLTY-----FPVADLSIIAALYARFTAQIRGAVDLSLYPREGV 112
Db 65 AAVAAGDISQATTGGNDIIMGNSVPPVVTHELIALGKFKQT-ILSVVE---KPKPHRQ 120

QY 113 SSRELKVKVSDVWNSLSRSYFKDRAHIQSLFSFITGKLDSSGVAFVAVGACQALGLRD 172
Db 121 ATREVTKVSDVWNSLSRSYFKDRAHLQNLYSYLSGNKLDLFCGVALAVVAGCOLLYGKD 180

QY 173 VHLAISEDHAWVYFGPNEGQTAETVTHGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDR 232
Db 181 VRLAISEDHAWVYFGQKRVETIEVTHGKGSDEKRGQDIRPGIESGSWLYLGLGLAVVCCR 240

QY 233 KMEVAFMVCAINPSIDLHTSDLSLELLQOQKLLWLLDGLHLERYPMALGNLADLEELEPT 292
Db 241 GMEVAACAALNLSLTSNSDCVEAELOQQLLWLLDGLHLKRYPMALGTLGEELEHRT 300

QY 293 Y-CREDEEYKFEFFVANDVIPNLLKEAASLLEAGEERPGESQSQSQSQ-----SALQDP 407
Db 361 YQCRDEEYKFEFFVANDVIPNLLKEAASLLEAGEERPGESQSQSQSQ-----SALQDP 401

QY 408 ECFHLLRFYDGICKWEGSPPTVLHVGWATFLVQSLGRFEGOVROKVIYVSREAAEAEEPEWCEAREGR 457
Db 408 ECFHLLRFYDGICKWEGSPPTVLHVGWATFLVQSLGRFEGOVROKVIYVSREAAEAEEPEWCEAREGR 457
```



```
Db 241 GMEVAALCAALNLSRNSDCVEAEVLEQQQLLWLLVDLGHLLKRYPMALGTLGELIEIHT 300
QY 293 PGRPDPLTLVHKGIASAKTYRDEHIYPYMYLAGYHCRNRNVRREALQAWADTATVIQDYN 352
Db 301 HPSISCEQLYREALIESARTHYRNHHVPTYQGYNRLKLYRDFAAFAANAADVIRLYT 360
QY 353 Y-CREDEEYKEFEFVANDVIPNLLKEAASLLEAGEERPGEGSQGTOSQ-----SALQDP 407
Db 361 YOCRDDEEYKELLDIANELIPIYMK-----TESSGHSARSILRDS 401
QY 408 ECFALLRFYDGIKCKEESGPTVLVHGWATFVLSLGRFEGOVROKVRT----- 457
Db 402 EVFANLLRFYDGIKCKEESGPTVLVHGWATFVLSLGRFEGOVROKVRT-----SALQDP 461
QY 458 -----VSREA-EAAAEAEPEWEEAREGRRR-----GPRRESKPEPP--- 493
Db 462 AKAEARAEQEAKEAESAAGSEAMEGNNRMTATKEERNVLAALGKELSLKPRNSQSSL 521
QY 494 -----GTGQ-----GAVSGPP-----PKKPAIDKGL----- 504
Db 522 LLINISVLTETTSYLLAVTLFTMLVHGLVSFVAESAMVLYKSKNSELPPTTLADLTAACG 581
QY 505 -----GTGQ-----GAVSGPP-----RKPPGTV 522
Db 582 EKILNDFLQGGQPFADQKQPSGSGESDNPELHNNNNNSNNNNNNHNAADKKEAAAT 641
QY 523 AGTARGEGGSTAQVPAPASPPPEG----- 548
Db 642 TTNATTTNSGSGTSVQLPVSEANNAQAGSQVQINDOLGKPKHKEAKKEETSDDYDPPE 701
QY 549 -----PVLTFQSEKMGKMKELLVATKINSIAIKLQTAQSQVOMKQKVSTPS 596
Db 702 IMLKRPVITLYSQMKRGLKDLLEAKLNTHAISLQVTAQSVASRKVRGVEKHS 754

RESULT 7
US-60-219-005-306
; Sequence 306, Application US/60219005
; GENERAL INFORMATION:
; APPLICANT: Skupski, Marion
; APPLICANT: et al.
; TITLE OF INVENTION: ISOLATED DROSOPHILA PROTEINS THAT SHOW
; TITLE OF INVENTION: HIGH HOMOLOGY TO HUMAN DISEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES ENCODING THESE DROSOPHILA PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL000734
; CURRENT APPLICATION NUMBER: US/60/219,005
; CURRENT FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/171,625
; PRIOR FILING DATE: 1999-12-23
; PRIOR APPLICATION NUMBER: 60/175,870
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/187,207
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: 60/191,639
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 342
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 306
; LENGTH: 813
; TYPE: PRT
; ORGANISM: Drosophila
US-60-219-005-306
```

```
Query Match 35.3%; Score 1124.5; DB 23; Length 813;
Best Local Similarity 35.4%; Pred. No. 3.4e-85;
Matches 274; Conservative 88; Mismatches 198; Indels 213; Gaps 16;

QY 9 TLPFLRSIDVWRLFAAEL-GREEPDVLVLSLVLFVGFVEHFLAVNRVPTNPVLTQPS 67
Db 10 SLFPLKSTADVINFLRALTSIGIEPDLTLLSIVVGVIYELSLTTGEAAQ-----AAQAAA 64
```

```
QY 68 APPPPGGLTY-----FPVADLSIIAALYARFTAOIRGAVIJSLYPPREGV 112
Db 65 AAAYAGDISQATTCGGNDIIMGNSVPFPVTHLITAGLYKKFQT-ILSWE---KPKPHRQ 120
QY 113 SSRELKVKSDVWNSLSRSFYFKDRAHIQSILFSFITGTLKLDSSGVAFVAVGACACGLRD 172
Db 121 ATREVTKKSDVWNSLSRSFYKDRALHQLNLYSYLSNKLKDCFGVALAVVAGCOLLYGKD 180
QY 173 VHLALSADHAWVFGPNGEQTAETVTHGKGNEDRRGTVMAGVAERSWLVIKGSYMRCDR 232
Db 181 VRLAISADHAWVFGQKRVETIEVTHGKGSDEKRGODIRPIBGSWLYLGLGLAVVCCR 240
QY 233 KMEVAFMVCAINPSIDLHSDLSLELQLOKLLWLLYDLGHLERYPMALGNLADLEELEPT 292
Db 241 GMEVAALCAALNLSRNSDCVEAEVLEQQQLLWLLVDLGHLLKRYPMALGTLGELIEIHT 300
QY 293 PGRPDPLTLVHKGIASAKTYRDEHIYPYMYLAGYHCRNRNVRREALQAWADTATVIQDYN 352
Db 301 HPSISCEQLYREALIESARTHYRNHHVPTYQGYNRLKLYRDFAAFAANAADVIRLYT 360
QY 353 Y-CREDEEYKEFEFVANDVIPNLLKEAASLLEAGEERPGEGSQGTOSQ-----SALQDP 407
Db 361 YOCRDDEEYKELLDIANELIPIYMK-----TESSGHSARSILRDS 401
QY 408 ECFALLRFYDGIKCKEESGPTVLVHGWATFVLSLGRFEGOVROKVRT----- 457
Db 402 EVFANLLRFYDGIKCKEESGPTVLVHGWATFVLSLGRFEGOVROKVRT-----SALQDP 461
QY 458 -----VSREA-EAAAEAEPEWEEAREGRRR-----GPRRESKPEPP--- 493
Db 462 AKAEARAEQEAKEAESAAGSEAMEGNNRMTATKEERNVLAALGKELSLKPRNSQSSL 521
QY 494 -----GTGQ-----GAVSGPP-----PKKPAIDKGL----- 504
Db 522 LLINISVLTETTSYLLAVTLFTMLVHGLVSFVAESAMVLYKSKNSELPPTTLADLTAACG 581
QY 505 -----GTGQ-----GAVSGPP-----RKPPGTV 522
Db 582 EKILNDFLQGGQPFADQKQPSGSGESDNPELHNNNNNSNNNNNNHNAADKKEAAAT 641
QY 523 AGTARGEGGSTAQVPAPASPPPEG----- 548
Db 642 TTNATTTNSGSGTSVQLPVSEANNAQAGSQVQINDOLGKPKHKEAKKEETSDDYDPPE 701
QY 549 -----PVLTFQSEKMGKMKELLVATKINSIAIKLQTAQSQVOMKQKVSTPS 596
Db 702 IMLKRPVITLYSQMKRGLKDLLEAKLNTHAISLQVTAQSVASRKVRGVEKHS 754

RESULT 8
PCT-US01-00663-31783
; Sequence 31783, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Rank, David R.
; APPLICANT: Hanzel, David K.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FO
; TITLE OF INVENTION: ANALYSIS OF GENE EXPRESSION IN HUMAN PLACENTA
; FILE REFERENCE: PB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; CURRENT FILING DATE: 2001-01-30
; PRIOR APPLICATION NUMBER: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263,6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
```

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1151 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-023-905-4

Query Match	4.1%	Score	131;	DB	14;	Length	1151;	
Best Local Similarity	21.4%;	Pred.	No.	0.41;				
Matches	118;	Conservative	73;	Mismatches	220;	Indels	140;	
Gaps	26;							
QY	119	KKVSDVWNSLSRYF	---	KDRAHIQSLFSGFTGT	KDDSGCAFVAVGACQALGRDVHL	175		
DB								
DB	382	KKCFDLI	-	SHNTYIFQADEQEF	VIWISVYNSKEALNMAFRGQSAGDSDLELTK	439		
QY	176	ALSEDAWVVFGNGEQTA	EVTHGKGNEDRQGT	VNAGVAERSWLYKGSYMRCDRME	235			
DB								
DB	440	AIITEDVLR	-----	PGNE	-----	VCCDCGVPEPKWLSFNLGILTC	474	
QY	236	VAFWVCA	-	INPSIDLH	---	TDSLELLOLQKLLWLYDLGLHRYPMALGNLADELEP	291	
DB								
DB	475	---	IECSGTHREMGVHIS	IQSMELDKLGTSELLAKNVGNSSNEILEGNL	524			
QY	292	TPC	-	RDPDLTYHKGASAKTYRD	---	EHYPYMYLAGYHCRNRNRYREALQAWADTATV	347	
DB								
DB	525	SPSPKAP	---	SSDTERKEYINAKYVHRFARTATTATARQDLYEAVRT	-	RDLMAL	579	
QY	348	IQDYNCREDEET	YKFEFFVANDVIPNL	-----	KEASLLEAGEERPGESQGTQS	399		
DB								
DB	580	IQLY	---	ADGVELMDPPEAGDPGETALHFAVRTSDQTS	SLHLVDFLVQNSGTTLDROTES	636		
QY	400	QGSAL	-----	QDPECAHLLRYVDGICKWEEGSP	---	VLVHGWATFLVQ	-	
DB								
DB	637	GNAALHYCCTYERKEPKU	LLLRKPKPSIDLNVQNETALDIARRUNVQCELLVEAAAGR	696				
QY	447	FEQVRQKVRIVSREAAEA	EPWGEAREGR	-	RGPRES	---	KPEPPPPKPPALDKG	503
DB								
DB	697	FNPVHVVEYEWNLRL	EEIDSDDDLDDK	PSVVKERSPRQSFCHSSSVSQEKLTPGY	756			
QY	504	LG	-----	TCQGVSGP	-----	PKPPG	-	
DB								
DB	757	LGHDRKQRLSYGAFANPV	YSTETPASVSEGTIAKTPAKAPSCGPPTSLPLGSOSS	816				
QY	542	---	ASPPPG	-----	PVL	-----	TFQSEKMGKME	563
DB								
DB	817	AGCSSTLSKKRAPPPPP	CHRTHSDPPSPVLOGPSKGSBTPPSANRTPSPANFEGIOQ	876				
QY	564	LLVATKINSSA	574					
DB								
DB	877	QOOSTSMNTRK	887					

```

RESULT 15
US-60-191-637-9944
US-60-191-637-9944
; Sequence 9944, Application US/60191637
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; TITLE OF INVENTION: NUCLEIC ACID DETECTION KITS COMPRISING
; TITLE OF INVENTION: GENES SEQUENCES EXPRESSED FROM THE DROSOPHILA GENOME, AND
; TITLE OF INVENTION: USES THEREOF
; FILE REFERENCE: CL000392
; CURRENT APPLICATION NUMBER: US/60/191,637
; CURRENT FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 42660
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9944
; LENGTH: 566
; TYPE: PRT
; ORGANISM: DROSOPHILA

```

US-60-191-637-9944

Query Match	4.1%	Score 129;	DB 23;	Length 566;
Best Local Similarity	25.4%;	Pred. No. 0.2;		
Matches	62;	Conservative	14;	Mismatches 90; Indels 78; Gaps 9;
QY	358	EEYKKEFFEVANDVIPNLLKEAASLL	EAGEERGEQSQGTQSGSALQD-----	406
		:	:	:
Db	266	EEYKNOIKITLN-----TRLKEA----	EAAEFAERSVKLQKEVDRLEDIDLVEKERYCM	317
		:	:	:
QY	407	-----PECEAHLLRFYDGICKWEBS-	-PTPVLHVGMATFLVQSLGRFEGQVQKVRIV	458
		:	:	:
Db	318	IGSLDBEAFVDLIKLEPF--WNPRNPKPTPKL----	PTTPEELAAEAEARAAASAAA	371
		:	:	:
QY	459	SRAEAAAEPEPWGEAEAREGRRRPRESKP-	EPPEPPPKPALDKGLGTGGVAGSGPPRK	517
		:	:	:
Db	372	AAEAEEAAANAAGEACPDGAAPGEEKAPAKE	PTPKPP-----TPPPPP	418
		:	:	:
QY	518	PPGTVACTARGPEGGTAQV-----	PAPAAASPPP	546
		:	:	:
Db	419	PPPFYSIDLPPGEAGPYVKNVEPPPPGSEPE	PVPAAGEGAAPAECAAPPAEAGAAPPA	478
		:	:	:
QY	547	EGPV	550	
		:	:	
Db	479	EGAV	482	
		:	:	

Search completed: October 12, 2001, 18:20:49
Job time: 207 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:14:02 ; Search time 13.8 seconds
(without alignments)
910.153 Million cell updates/sec

Title: US-09-380-337-2
Perfect score: 3185
Sequence: 1 MGLKAAQKTLFPLRSIDVV.....KVSTPDSYTLFLKQRKGL 610

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 197339 seqs, 20590346 residues

Total number of hits satisfying chosen parameters: 197339

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/2/1aa/5A_COMB.pep:*
2: /cgn2_6/ptodata/2/1aa/5B_COMB.pep:*
3: /cgn2_6/ptodata/2/1aa/6A_COMB.pep:*
4: /cgn2_6/ptodata/2/1aa/6B_COMB.pep:*
5: /cgn2_6/ptodata/2/1aa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/2/1aa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3185	100.0	610	2 US-08-865-337A-6	Sequence 6, Appli
2	2778	87.2	564	2 US-08-865-337A-1	Sequence 1, Appli
3	117.5	3.7	124	2 US-08-925-237-2	Sequence 2, Appli
4	117.5	3.7	1581	4 US-09-110-517-2	Sequence 2, Appli
5	117	3.7	309	2 US-08-405-175A-7	Sequence 7, Appli
6	117	3.7	696	3 US-08-906-865-4	Sequence 4, Appli
7	114.5	3.6	535	4 US-09-029-348-1	Sequence 1, Appli
8	114.5	3.6	537	4 US-09-029-348-4	Sequence 4, Appli
9	108.5	3.4	2293	4 US-09-368-590-2	Sequence 2, Appli
10	108	3.4	2205	1 US-08-093-453B-2	Sequence 2, Appli
11	107.5	3.4	1321	2 US-08-317-310A-64	Sequence 64, Appli
12	106.5	3.3	309	2 US-08-405-175A-8	Sequence 8, Appli
13	106	3.3	928	5 PCT-US94-10357-3	Sequence 3, Appli
14	105.5	3.3	82	2 US-08-925-237-4	Sequence 4, Appli
15	105	3.3	501	2 US-08-969-630-4	Sequence 4, Appli
16	105	3.3	816	1 US-08-038-760-3	Sequence 3, Appli
17	105	3.3	816	2 US-08-470-091-3	Sequence 3, Appli
18	105	3.3	928	1 US-08-204-329-1	Sequence 1, Appli
19	105	3.3	928	2 US-08-959-638-8	Sequence 8, Appli
20	105	3.3	928	2 US-08-482-627-5	Sequence 5, Appli
21	105	3.3	928	3 US-08-801-092-4	Sequence 4, Appli
22	105	3.3	928	4 US-08-328-673A-8	Sequence 4, Appli
23	105	3.3	1009	2 US-08-357-642A-1	Sequence 1, Appli
24	105	3.3	1009	2 US-08-460-626-1	Sequence 1, Appli
25	104	3.3	502	2 US-08-969-630-2	Sequence 2, Appli
26	104	3.3	534	4 US-09-029-348-5	Sequence 5, Appli
27	104	3.3	661	2 US-08-795-868-14	Sequence 14, Appli

28	104	3.3	928	5 PCT-US94-10357-2	Sequence 2, Appli
29	103	3.2	1149	3 US-08-560-005-5	Sequence 5, Appli
30	103	3.2	1479	3 US-08-840-062-2	Sequence 2, Appli
31	101.5	3.2	1418	3 US-08-963-825-20	Sequence 20, Appli
32	101.5	3.2	1418	4 US-09-010-999-1	Sequence 1, Appli
33	100.5	3.2	478	3 US-08-155-888-2	Sequence 2, Appli
34	100.5	3.2	1366	3 US-08-963-825-19	Sequence 19, Appli
35	100	3.1	1479	3 US-08-840-062-4	Sequence 4, Appli
36	99.5	3.1	565	4 US-08-961-083-218	Sequence 218, Appli
37	99.5	3.1	866	1 US-08-386-727-8	Sequence 8, Appli
38	99.5	3.1	866	2 US-08-600-452A-8	Sequence 8, Appli
39	99.5	3.1	1805	1 US-07-853-913-2	Sequence 2, Appli
40	99	3.1	1239	2 US-08-937-931-2	Sequence 2, Appli
41	99	3.1	1239	4 US-09-285-502-2	Sequence 2, Appli
42	98.5	3.1	258	4 US-08-961-083-90	Sequence 90, Appli
43	98.5	3.1	267	1 US-07-928-611-13	Sequence 13, Appli
44	98.5	3.1	267	2 US-08-487-811A-13	Sequence 13, Appli
45	98.5	3.1	267	5 PCT-US93-07370-13	Sequence 13, Appli

ALIGNMENTS

RESULT 1
US-08-865-337A-6
; Sequence 6, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 610 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GenBank
; CLONE: 1945387
US-08-865-337A-6

```
Query Match      100.0%; Score 3185; DB 2; Length 610;
Best Local Similarity 100.0%; Pred. No. 1.1e-293;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MGLKAAQKTLFPLRSIDDDVVRFLFAAELGREPDLVLLSLVLFGEVHEFLAVNRVPTNYPE 60
DB 1 MGLKAAQKTLFPLRSIDDDVVRFLFAAELGREPDLVLLSLVLFGEVHEFLAVNRVPTNYPE 60

QY 61 LTFQSPAPDPGGTLTFPPVADLSIIAALYARFQAIRGAVDLSLYPREGVSSRELKVK 120
DB 61 LTFQSPAPDPGGTLTFPPVADLSIIAALYARFQAIRGAVDLSLYPREGVSSRELKVK 120

QY 121 VSDVIWNSLSRSYFKDRAHIQSLFSFTGTGLDSSGVAFVAVGACQALGRDVLHALSED 180
DB 121 VSDVIWNSLSRSYFKDRAHIQSLFSFTGTGLDSSGVAFVAVGACQALGRDVLHALSED 180

QY 181 HAWVVFPGNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240
DB 181 HAWVVFPGNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240

QY 241 CAINPSIDLHTDLSLELLQLOQKLLWLLYDLGLHLERYPMALGNLADLELEPTGCRPDPLT 300
DB 241 CAINPSIDLHTDLSLELLQLOQKLLWLLYDLGLHLERYPMALGNLADLELEPTGCRPDPLT 300

QY 301 LYHKGIASAKTYRDEHIYPYMYLAGYHCRNRNVRREALQAWADTATVIQDYNCREDEEI 360
DB 301 LYHKGIASAKTYRDEHIYPYMYLAGYHCRNRNVRREALQAWADTATVIQDYNCREDEEI 360

QY 361 YKEFEFVANDVIPNLLKEAASLLEAGEERPEQSGQTSQGSALQDPECFAHLLRFYDGI 420
DB 361 YKEFEFVANDVIPNLLKEAASLLEAGEERPEQSGQTSQGSALQDPECFAHLLRFYDGI 420

QY 421 CKWEGSPTPLVHVGWATFLVQSLGRPEGOVRQKRVTVSREAEAAAEPEWGEAREGR 480
DB 421 CKWEGSPTPLVHVGWATFLVQSLGRPEGOVRQKRVTVSREAEAAAEPEWGEAREGR 480

QY 481 RGPRESKPEEPPPPKPPALDKGLGTGOGAVSGPPRPKPGTAVAGTARGPEGGSTAQVPAP 540
DB 481 RGPRESKPEEPPPPKPPALDKGLGTGOGAVSGPPRPKPGTAVAGTARGPEGGSTAQVPAP 540

QY 541 AASPPPEGPVLTFOSEKMKGMKELLVATKINSSAIKQLTAQSOVQMKQKVSTPSPDYTL 600
DB 541 AASPPPEGPVLTFOSEKMKGMKELLVATKINSSAIKQLTAQSOVQMKQKVSTPSPDYTL 600

QY 601 SFLKRQRKGL 610
DB 601 SFLKRQRKGL 610
```

```
RESULT 2
US-08-865-337A-1
; Sequence 1, Application US/08865337A
; Patent No. 5975649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murty, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESS: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
```

```
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 564 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT14
; CLONE: Consensus
; US-08-865-337A-1
```

```
Query Match      87.2%; Score 2778; DB 2; Length 564;
Best Local Similarity 89.5%; Pred. No. 4.2e-255;
Matches 547; Conservative 2; Mismatches 14; Indels 48; Gaps 4;

QY 1 MGLKAAQKTLFPLRSIDDDVVRFLFAAELGREPDLVLLSLVLFGEVHEFLAVNRVPTNYPE 60
DB 1 MGLKAAQKTLFPLRSIDDDVVRFLFAAELGREPDLVLLSLVLFGEVHEFLAVNRVPTNYPE 60

QY 61 LTFQSPAPDPGGTLTFPPVADLSIIAALYARFQAIRGAVDLSLYPREGVSSRELKVK 120
DB 61 LTFQSPAPDPGGTLTFPPVADLSIIAALYARFQAIRGAVDLSLYPREGVSSRELKVK 120

QY 121 VSDVIWNSLSRSYFKDRAHIQSLFSFTGTGLDSSGVAFVAVGACQALGRDVLHALSED 180
DB 121 VSDVIWNSLSRSYFKDRAHIQSLFSFTGTGLDSSGVAFVAVGACQALGRDVLHALSED 180

QY 181 HAWVVFPGNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240
DB 181 HAWVVFPGNGEQTAEVTHWKGKGNEDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFMV 240

QY 241 CAINPSIDLHTDLSLELLQLOQKLLWLLYDLGLHLERYPMALGNLADLELEPTGCRPDPLT 300
DB 241 CAINPSIDLHTDLSLELLQLOQKLLWLLYDLGLHLERYPMALGNLADLELEPTGCRPDPLT 300

QY 301 LYHKGIASAKTYRDEHIYPYMYLAGYHCRNRNVRREALQAWADTATVIQDYNCREDEEI 359
DB 301 LYHKGIASAKTYRDEHIYPYMYLAGYHCRNRNVRREALQAWADTATVIQDYNCREDEEI 359

QY 361 YKEFEFVANDVIPNLLKEAASLLEAGEERPEQSGQTSQGSALQDPECFAHLLRFYDGI 419
DB 361 YKEFEFVANDVIPNLLKEAASLLEAGEERPEQSGQTSQGSALQDPECFAHLLRFYDGI 419

QY 420 CKWEGSPTPLVHVGWATFLVQSLGRPEGOVRQKRVTVSREAEAAAEPEWGEAREGR 479
DB 420 CKWEGSPTPLVHVGWATFLVQSLGRPEGOVRQKRVTVSREAEAAAEPEWGEAREGR 479

QY 480 RGPRESKPEEPPPPKPPALDKGLGTGOGAVSGPPRPKPGTAVAGTARGPEGGSTAQVPAP 539
DB 480 RGPRESKPEEPPPPKPPALDKGLGTGOGAVSGPPRPKPGTAVAGTARGPEGGSTAQVPAP 539

QY 540 PAASPPPEGPVLTFOSEKMKGMKELLVATKINSSAIKQLTAQSOVQMKQKVSTPSPDYT 599
DB 540 PAASPPPEGPVLTFOSEKMKGMKELLVATKINSSAIKQLTAQSOVQMKQKVSTPSPDYT 599

QY 600 LSFLKRQRKGL 610
```



```
Db 554 LSFLKQAKGL 564
||||| |||
RESULT 3
US-08-925-237-2
; Sequence 2, Application US/08925237
; Patent No. 5981720
; GENERAL INFORMATION:
; APPLICANT: Azen, Edwin A.
; APPLICANT: Pan, David
; TITLE OF INVENTION: Human Salivary Proteins And Fragments
; TITLE OF INVENTION: Thereof Having Alpha-Glucosidase Inhibitory Activity
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kent Barta
; STREET: 100 East Wisconsin Avenue
; CITY: Milwaukee
; STATE: WI
; COUNTRY: USA
; ZIP: 53202-4108
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: PatentIn Ver. 1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/925,237
; FILING DATE:
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Barta, Kent S.
; REGISTRATION NUMBER: 29,042
; REFERENCE/DOCKET NUMBER: 96429/9002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (608) 257-3501
; TELEFAX: (608) 283-2275
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 124 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-925-237-2

Query Match 3.7%; Score 117.5; DB 2; Length 124;
Best Local Similarity 32.3%; Pred. No. 0.0013;
Matches 31; Conservative 9; Mismatches 25; Indels 31; Gaps 5;

QY 469 EPWGEAREGRRGRPRE--SKPEEPPPPKPKALDKGLGTGOGAVSGPP-----R 516
Db 16 QPQGPVPPGKPGPPQGGNKKQGPVPPGKPP-----QGPPPGDNKQSAR 62
QY 517 KPQGTVAGTARGEGGSTAQVAP-----AASPPPEG 548
Db 63 SPQKPGQGP--PQGGNQPGPPPPGKPGQPPPG 96

RESULT 4
US-09-110-517-2
; Sequence 2, Application US/09110517A
; Patent No. 6248520
; GENERAL INFORMATION:
; APPLICANT: Roeder, Robert G
; APPLICANT: Fondell, Joseph D
; APPLICANT: Yuan, Chao X
; APPLICANT: Ito, Mitsuhiro
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING NUCLEAR HORMONE
; TITLE OF INVENTION: RECEPTOR COACTIVATORS AND USES THEREOF
; FILE REFERENCE: 600-1-224
; CURRENT APPLICATION NUMBER: US/09/110,517A
; CURRENT FILING DATE: 1998-07-06
```

```
; NUMBER OF SEQ ID NOS: 51
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 1581
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-110-517-2

Query Match 3.7%; Score 117.5; DB 4; Length 1581;
Best Local Similarity 22.4%; Pred. No. 0.075;
Matches 72; Conservative 33; Mismatches 128; Indels 89; Gaps 13;

QY 341 WADTATVIQDNYNCREDEEYKFEFVANDVIPNLLKEAASLLLEAGEERPEQSQ----- 395
Db 835 YTDPADLIADAAGSPSSDSPTNHFFHGDVDFNPDLN-NSQSQSGFGEEYDESSQSGDND 893
QY 396 ----GTQSQ-----GSALQDPECFALHLLRF 416
Db 894 DFKGFASQALNTLGVPMLGDNGETKFKGNQADTVDFSIISVAGKALAPADLMEH----- 949
QY 417 YDGICKWEESPTVPLHVGHATFVQSLGRFEGOVROKVRIVSREAAEAEEPPWGEAR 476
Db 950 -----HSGSQGPLLTG-----DLGKEKTQRRVK---EGNGTNSNTLSGP-GLDSK 991
QY 477 EGRR-RGPRRESKPEEPPPKPALDKGLGTGOGAVSGPPKPPGTVAGTARGEGGSTA 535
Db 992 PGKSRTPSNDGSKDKPPKPKKADTEGKSPSHSS-SNRFTTP-TSTGSGSKSPCSAGRS 1049
QY 536 QVPAPAAAPP-----PEGPVLTTFOSERKMGMKMKELLVATKINSSAIK-----LQL 579
Db 1050 QTPPGVATPPPKITTIQIPKGTVMV---GKPSHSQYTSSTSGSVSSSGSKSHSHSSSSSS 1106
QY 580 TAQSQVQMKKQKVSTPSTDYTLS 601
Db 1107 SASTSGMKMSKSESGSSSKLS 1128

RESULT 5
US-08-405-175A-7
; Sequence 7, Application US/08405175A
; Patent No. 5885772
; GENERAL INFORMATION:
; APPLICANT: Aderem, Alan A.
; APPLICANT: Chen, Jianmin
; APPLICANT: Chang, Sandy
; TITLE OF INVENTION: METHOD FOR THE DETECTION OF ANENCEPHALY
; NUMBER OF SEQUENCES: 12
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Klauber & Jackson
; STREET: 411 Hackensack Avenue
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/405,175A
; FILING DATE: 16-MAR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-121A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201 487-5800
; TELEFAX: 201 343-1684
; TELEX: 133521
; INFORMATION FOR SEQ ID NO: 7:
```

; SEQUENCE CHARACTERISTICS:
; LENGTH: 309 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; DESCRIPTION: predicted primary structure of murine MARCKS
; HYPOTHETICAL: NO
US-08-405-175A-7

Query Match 3.7%; Score 117; DB 2; Length 309;
Best Local Similarity 26.3%; Pred. No. 0.0062;
Matches 54; Conservative 21; Mismatches 86; Indels 44; Gaps 8;
QY 377 KEASLLEAEEERQEQSQGTQSGSALODPECFHLLRLFYDVGICKWEEGSPPTVLHWG 436
Db 101 KEAAEAEPAPESPAAEAGASASTSPKAE-----DGAAP-SPSETPKKKKKR 150
QY 437 ATFLVQ-SLGRFEGQVRQKVRIVSREAAEAE-----EPWGEAR 476
Db 151 FSKKSKLGSFSSKSKESGEAGAEAGATAEGAKDEAAAAAGGEGAAAPGEAGAGA 210
QY 477 EGRRRGPRRE---SKPEPPPKKPKALDKGLGTGQAVSGPPRK--PPGTVAG---TARG 528
Db 211 EGAAGGEPREAAEAPEQPEQPEQAAEPPQAEQSEAAAGEAEPAPGATAGDASSAAG 270
QY 529 PEGGTAQV-----PAPAAAPPE 547
Db 271 PEQAPANTDEAAASAPAAPEPQ 295

RESULT 6
US-08-906-865-4
; Sequence 4, Application US/08906865
; Patent No. 6040168
; GENERAL INFORMATION:
; APPLICANT: Greengard, Paul
; APPLICANT: Porton, Barbara
; APPLICANT: Kao, Hung-Teh
; TITLE OF INVENTION: DNA ENCODING THE HUMAN SYNAPSIN III GENE
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: David A. Jackson, Esq.
; STREET: 411 Hackensack Ave, Continental Plaza, 4th
; CITY: Hackensack
; STATE: New Jersey
; COUNTRY: USA
; ZIP: 07601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/906.865
; FILING DATE:
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Jackson Esq., David A.
; REGISTRATION NUMBER: 26,742
; REFERENCE/DOCKET NUMBER: 600-1-202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 201-487-5800
; TELEFAX: 201-343-1884
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 696 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; MOLECULE TYPE: protein
; DESCRIPTION: /desc - "Synapsin Ia"
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapiens
US-08-906-865-4

Query Match 3.7%; Score 117; DB 3; Length 696;
Best Local Similarity 20.5%; Pred. No. 0.023;
Matches 130; Conservative 68; Mismatches 214; Indels 222; Gaps 34;
QY 66 SPADPPPGGLTYFFVADLSI---IAALYARETAQIRGAVDLSLYPREGGVSSRELV---- 118
Db 65 APSGSGGGGGFFSSLSNAVKQTAAAAATSEQVGGG---SGGAGRGGAASRLVLVIDE 121
QY 119 -----KKVSDVIWNSLSRSYFKDR---AHQSLFSFTGTCLDSSGVAFAVUGA 164
Db 122 PHTDWAKYFGKKIKGEIDIKVEQAEFSDLNLVAHANGFS--VDMVLRNGV--KVYRS 177
QY 165 COALGLRDVHLALSSEDHAWVVFGPNGEQTAEVTHWKGNEDERRGOTVN---AGVAERSML 221
Db 178 LKP-----DEVLIHQAFSM-----ARNGDYSLVICLQVAGIPSVNSL 216
QY 222 YLKGSMRCDRKMEVAFMVCAINPSIDLH-----TDSLELL-----OLQOKLLWL 267
Db 217 H--SVYNFCDRKPMVFAOMV-----RLHKKLGTETFPPLIDQTFVPHKHEMLSTTYPVV 267
QY 268 YDLGHLE--RYPMALGNLADLEELEPTGPRDPLTLYHKGIASAKTYRDEHIYMYLA 325
Db 268 VKMGHAGSMGKVKVDNQHDFDIASV-----VALTKTYATAE---PFID-A 310
QY 326 GYHCRNRNREALQA-----WADTATV 348
Db 311 KYDVRVQKIGONYKAYMRTSVGNWKTNTGSAMLEQIAMSDBRYKLWVDTCSEIFEGGLDIC 370
QY 349 -----QDY-----NYCREDEITYKEFFEVAN---DVTINLKEAAS- 381
Db 371 AVEALHGKDRDHIEVVGSSMPLIGDHQDEKQLIVEL--VWNKMAQALPRQRORDASP 428
QY 382 -----LLEAGEER-----PGEOSGTOSQGSALQD---PE 408
Db 429 GRGSHGOTPSGALPLGKQTSSQAGPPAQOEPPPPGPGPQPGPQPGPQPGPQ 488
QY 409 CFAHLRFYDGICKWEEGSPPTVLHVHWATFLVQSLGR---FEGQVRQKVRIVSREAEA 464
Db 489 GQOHL-----SGLGP-PAGSPLPQLPSPTSAPQOPASQAAPPTQCGGRQS-RPVAGGPGA 542
QY 465 AEAEPPWGEAEAREGRRGP---RRESKPEEPPPPKKPALDKGLGTGCGAVSGPPPKPGT 521
Db 543 PPAARPPASPSQ-RQAGPPQATRTQTSVSGPAPPKA-----SGAPPGQQRQGPQKPPGP 597
QY 522 VAGTARGPEGGSTAQVPAPAAASP-----PEGP 549
Db 598 APTRQASQAG-----PVPRTGPTTQPPRPSGP 626

RESULT 7
US-09-029-348-1
; Sequence 1, Application US/09029348
; Patent No. 6171827
; GENERAL INFORMATION:
; APPLICANT: THE VICTORIA UNIVERSITY OF MANCHESTER
; TITLE OF INVENTION: NOVEL PROCOLLAGENS
; FILE REFERENCE: G087857PUS LISTING
; CURRENT APPLICATION NUMBER: US/09/029.348
; CURRENT FILING DATE: 1998-05-07
; NUMBER OF SEQ ID NOS: 20
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 535
; TYPE: PRT
; ORGANISM: Artificial Sequence

FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: SEQUENCE
OTHER INFORMATION: DERIVED FROM CDNA OF PROCOLLAGENS
US-09-029-348-1

Query Match 3.6%; Score 114.5; DB 4; Length 535;
Best Local Similarity 34.2%; Pred. No. 0.026;
Matches 39; Conservative 5; Mismatches 31; Indels 39; Gaps 7;

QY 473 EEAREGRR--RGPRRESKPEEP-----PPKKPAL-----DKGLG 505
Db 28 ETVRKGPAGDRGRGRGPPGPRDGEDGPTGPPGPPGPGGFAAQYDGGVG 87
QY 506 TGQGVAYS--GPPRKPPTGVTAGTARGPEG-----GSTAQVPAPAAASPPPEGPV 550
Db 88 LGPGPGLMGPRGPPG--AAGAPGPGFQGPAGEGPEGQT--GPAGAPGHGPV 138

RESULT 8
US-09-029-348-4
Sequence 4, Application US/09029348
Patent No. 6171827
GENERAL INFORMATION:
APPLICANT: THE VICTORIA UNIVERSITY OF MANCHESTER
TITLE OF INVENTION: NOVEL PROCOLLAGENS
FILE REFERENCE: G087857PUS LISTING
CURRENT APPLICATION NUMBER: US/09/029,348
CURRENT FILING DATE: 1998-05-07
NUMBER OF SEQ ID NOS: 20
SOFTWARE: Patent In Ver. 2.0
SEQ ID NO 4
LENGTH: 537
TYPE: PRT
ORGANISM: Artificial Sequence
OTHER INFORMATION: Description of Artificial Sequence: SEQUENCE
OTHER INFORMATION: DERIVED FROM CDNA OF PROCOLLAGENS
US-09-029-348-4

Query Match 3.6%; Score 114.5; DB 4; Length 537;
Best Local Similarity 34.2%; Pred. No. 0.026;
Matches 39; Conservative 5; Mismatches 31; Indels 39; Gaps 7;

QY 473 EEAREGRR--RGPRRESKPEEP-----PPKKPAL-----DKGLG 505
Db 28 ETVRKGPAGDRGRGRGPPGPRDGEDGPTGPPGPPGPGGFAAQYDGGVG 87
QY 506 TGQGVAYS--GPPRKPPTGVTAGTARGPEG-----GSTAQVPAPAAASPPPEGPV 550
Db 88 LGPGPGLMGPRGPPG--AAGAPGPGFQGPAGEGPEGQT--GPAGAPGHGPV 138

RESULT 9
US-09-368-590-2
Sequence 2, Application US/09368590
Patent No. 6187563
GENERAL INFORMATION:
APPLICANT: Solimena, Michele
TITLE OF INVENTION: INTERACTING POLYPEPTIDES FOR
FILE REFERENCE: 101918-200 (OCR-941)
CURRENT APPLICATION NUMBER: US/09/368,590
CURRENT FILING DATE: 1999-08-04
EARLIER APPLICATION NUMBER: 60/095,657
EARLIER FILING DATE: 1998-08-07
NUMBER OF SEQ ID NOS: 8
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 2
LENGTH: 2293
TYPE: PRT
ORGANISM: Human

US-09-368-590-2

Query Match 3.4%; Score 108.5; DB 4; Length 2293;
Best Local Similarity 21.0%; Pred. No. 0.98;
Matches 92; Conservative 75; Mismatches 153; Indels 119; Gaps 25;

QY 168 LGLRDVHLALSEDH-----AWVFGPNGEQTAEVTHGKGNEDRRGQTVNAGVA---ERS 219
Db 1336 LGFQEL-LMSEDKGVKVRPSWCGGPG-----GAGAGGRCRLIVGALCPQDEQS 1384
QY 220 WLYLKSVMRCDRKMEVAFVVCAINPSIDLHTSLELLQQLLLWLLYDLGLHLYPMA 279
Db 1385 TLQLLKKHLQLEQ-----GVENEYSIAQLSRQCR---ALLEMGH-----1421
QY 280 LGNLADLELEPTGPRDPLTLVYHKGITASAKTYRDEHIYPMYLAGYHCRNVRNALQ 339
Db 1422 ----PQSEIISRQSQVDRILYVALKELGEERRVLEQOYWLQY-----SROYSE-LE 1469
QY 340 AWADTATVI-----QDYNCREDEEIIYKEFFEVDVNDVIPNLLKEAASLLEAGEERPE 392
Db 1470 HWIAKEVYVAGSPGLGDFEHVSVLQEKFSEF-----ASETCMAGRERLAA 1515
QY 393 QSGTQSQGSALQDP--ECFAHLLRFYDGICKWEESPTPLVHVGMATFLVQSLGRFEG 450
Db 1516 VND-----MVDELIEC-GHTAA--ATWAEWKDG-----LNEAWAE-LLELMG-----1553
QY 451 VROKVRIVSREAEA--AEAEPPNGEAREGRRRGRPRRESKPEPPPPKPKALDKGLGTQ 508
Db 1554 TRAQLLAASRELHKFTSDARELQG-QIEEKRRRLPLRTTTPPE--PRPSASSMORTLRAFE 1610
QY 509 GAVSGPPRPPGTGVTAGTARGPEGGSTAQVPAPAAASPPPEGPVLTFSQEKMKMKELLVAT 568
Db 1611 HDLQ-----LLVSQVROLQEGA-AQLRTVYAGEHAEA-IASREQEVLOGMKELLASAC 1660
QY 569 K-----INSSAIKLQLTQAQ 582
Db 1661 EDARLHVSSSTADALRFHSQ 1679

RESULT 10
US-08-093-453B-2
Sequence 2, Application US/08093453B
Patent No. 5439814
GENERAL INFORMATION:
APPLICANT: Frey, Teryl K.
APPLICANT: Dominguez, Geraldina
APPLICANT: Wang, Chin-Yen
TITLE OF INVENTION: Modified Infectious Rubella Virus
NUMBER OF SEQUENCES: 13
CORRESPONDENCE ADDRESS:
ADDRESSEE: Jamie L. Greene, Jones & Askew
STREET: 191 Peachtree Street, 37th Floor
CITY: Atlanta
STATE: Georgia
COUNTRY: United States
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50
COMPUTER: Macintosh
OPERATING SYSTEM: 7.0
SOFTWARE: Microsoft Word
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/093,453B
FILING DATE: 19 JUL 1993
CLASSIFICATION: 435
PRIOR APPLICATION NUMBER:
APPLICATION NUMBER: U.S. 07/722,334
FILING DATE: 28 JUN 1991
ATTORNEY/AGENT INFORMATION:
NAME: Greene, Jamie L.
REGISTRATION NUMBER: 32,467
REFERENCE/DOCKET NUMBER: 07362-0101

STATE: WI
COUNTRY: USA
ZIP: 53202-4108
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/925,237
FILING DATE:
CLASSIFICATION: 514
ATTORNEY/AGENT INFORMATION:
NAME: Barta, Kent S.
REGISTRATION NUMBER: 29,042
REFERENCE/DOCKET NUMBER: 96429/9002
TELEPHONE: (608) 257-3501
TELEFAX: (608) 283-2275
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 82 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-925-237-4

Query Match 3.3%; Score 105.5; DB 2; Length 82;
Best Local Similarity 30.8%; Pred. No. 0.0092;
Matches 26; Conservative 10; Mismatches 18; Indels 31; Gaps 4;
QY 469 EPGEEAREGRRRGRPRE--SKPEEPPPPKPKALDKLGTGGGAVSGPPRKPPTVAGTA 526
DB 16 QPGPPPPKPKQPGPPGKNGKPGPPPKP-----QGPP----- 51
QY 527 RGPPEGSTAQVPAPAA---SPPEG 548
DB 52 --PQGNKSQGPppppgkpgpppg 74

RESULT 15
US-08-969-630-4
Sequence 4, Application US/08969630A
Patent No. 5981248
GENERAL INFORMATION:
APPLICANT: Xu, Hua
TITLE OF INVENTION: MAMMALIAN CELL DEATH PREVENTING KINASE, DPK
FILE REFERENCE: A463
CURRENT APPLICATION NUMBER: US/08/969,630A
CURRENT FILING DATE: 1997-11-13
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 4
LENGTH: 501
TYPE: PRT
ORGANISM: mouse
US-08-969-630-4

Query Match 3.3%; Score 105; DB 2; Length 501;
Best Local Similarity 17.3%; Pred. No. 0.18;
Matches 92; Conservative 79; Mismatches 170; Indels 190; Gaps 23;
QY 151 KLDSGVAFVAVGACQ-----ALGLRDVHLALSEDHAWVFGPNGEQTAEVTW 198
DB 53 QIEGTGIS---MSACKETALLRELNHPNVIATLQKFLSHSDRKVWLLF---DYAKHDLW 105
QY 199 HGKGNDRRGQTAVAGVARSWLKGSYMRCDRMEVAFMVCAINPSIDLHTDSLELLQ 258
DB 106 H-----IINFHASKA---NKKPMQPKSMVKSLLYQLDGIHYLHAN----- 145
QY 259 LQQLKLWLLY-----DLGHLERYPMALGNLADLELEPTPGR 295

DB 146 -----WVLRDLKPANILVMGEGPERGRVKIADMGFARLNSPLKPLADLD----- 191
QY 296 PDPLTLVHKG---IASAKTYRDEHIYPMYLAG-----YHCRRNRNVK----- 335
DB 192 PVVTFWYRAPELLLGARHYTKAIDIWAIGCIFAELLTSEPIFHCROEDIKTSNPFHIDQ 251
QY 336 -----BALQAWADTATVIQDY---NYCREDEEYIKKEFFEYANDVIFPNLL 376
DB 252 LDRIFVMGFPADKDWEDIRKMPYPTLQKDFRRTTVANSSLIKYMEKUKVKPDSKVPELL 311
QY 377 KEASLLEAGEERPEGQSQTQSOGSALQDPECFALHLLRFDGICKWDEGSPTPVLHVGW 436
DB 312 LQKLLTMDPTKRITSEQ-----ALQDP-----YFOEDPLTLQV-- 345
QY 437 ATFLVQSLGRFEGQVROKRVIVSREAEAEAEPEWGEAREGRRRRGPRRESKPEEPPPK 496
DB 346 -----FAG-----CQIPYKREFLNEDEP--EE--KGDNQOQQQNPHQQPAAPA 386
QY 497 KPALDKGLGTGGGAVSGPPRKP-----GTVAGTARGPEGST-----AQVPAPAPAS 543
DB 387 Q-----QTAAPPQAPPQSSAQTNGTAGGATAGGGGAGAGAGLQHSQDPGLNQV 434
QY 544 PPPEGPVLTFQSEKMKMGKELLVATKINSSAIKQLTAQSOVQVMKKOKVST 594
DB 435 PPNKKPRIGPSGANS GG--PVMPSDYQHSSS---RLNYQSSVQSGSSQSQT 480

Search completed: October 12, 2001, 18:18:47
Job time: 285 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 12, 2001, 18:13:42 ; Search time 22.09 Seconds
(without alignments)
1674.089 Million cell updates/sec

Title: US-09-380-337-2

Perfect score: 3185

Sequence: 1 MGLKAAQKTLFLRLSIDDVV.....KVSTPSTDTLSFLKQRKGL 610

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 412676 seqs, 60623988 residues

Total number of hits satisfying chosen parameters: 412676

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_0601.*
1: /SIDS8/gcgdata/geneseq/geneseq/AA1980.DAT.*
2: /SIDS8/gcgdata/geneseq/geneseq/AA1981.DAT.*
3: /SIDS8/gcgdata/geneseq/geneseq/AA1982.DAT.*
4: /SIDS8/gcgdata/geneseq/geneseq/AA1983.DAT.*
5: /SIDS8/gcgdata/geneseq/geneseq/AA1984.DAT.*
6: /SIDS8/gcgdata/geneseq/geneseq/AA1985.DAT.*
7: /SIDS8/gcgdata/geneseq/geneseq/AA1986.DAT.*
8: /SIDS8/gcgdata/geneseq/geneseq/AA1987.DAT.*
9: /SIDS8/gcgdata/geneseq/geneseq/AA1988.DAT.*
10: /SIDS8/gcgdata/geneseq/geneseq/AA1989.DAT.*
11: /SIDS8/gcgdata/geneseq/geneseq/AA1990.DAT.*
12: /SIDS8/gcgdata/geneseq/geneseq/AA1991.DAT.*
13: /SIDS8/gcgdata/geneseq/geneseq/AA1992.DAT.*
14: /SIDS8/gcgdata/geneseq/geneseq/AA1993.DAT.*
15: /SIDS8/gcgdata/geneseq/geneseq/AA1994.DAT.*
16: /SIDS8/gcgdata/geneseq/geneseq/AA1995.DAT.*
17: /SIDS8/gcgdata/geneseq/geneseq/AA1996.DAT.*
18: /SIDS8/gcgdata/geneseq/geneseq/AA1997.DAT.*
19: /SIDS8/gcgdata/geneseq/geneseq/AA1998.DAT.*
20: /SIDS8/gcgdata/geneseq/geneseq/AA1999.DAT.*
21: /SIDS8/gcgdata/geneseq/geneseq/AA2000.DAT.*
22: /SIDS8/gcgdata/geneseq/geneseq/AA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3185	100.0	610	19 AAW29749	Tumor suppressor p
2	2778	87.2	566	20 AAW86348	Human multiple end
3	131	4.1	1151	19 AAW77287	Zebrafish differ
4	126	4.0	580	22 AAB60493	Human cell cycle a
5	120.5	3.8	1318	22 AAB26874	HSV-2 immediate ea
6	118.5	3.7	1583	21 AAB29662	Human tyrosine pho
7	118.5	3.7	1636	21 AAB29661	Human histidine do
8	118	3.7	1301	20 AAW92296	Human alpha-1 (XVI
9	118	3.7	1336	20 AAW90894	Human collagen 18
10	117.5	3.7	124	19 AAW50192	Amino acid sequenc
11	117.5	3.7	1566	20 AAY06309	Human p53 regulato

12	117.5	3.7	1581	21 AAY69669	Human thyroid rece
13	117.5	3.7	1581	21 AAY69671	Human thyroid rece
14	117.5	3.7	1581	21 AAY69672	Human thyroid rece
15	116	3.6	2971	21 AAB41231	Human OREX ORF995
16	116	3.6	2972	22 AAB50363	Human SRCAP. Homo
17	116	3.6	3118	22 AAB50362	Human muscarinic a
18	114.5	3.6	479	10 AAW92606	Truncated pro-alpha
19	114.5	3.6	535	18 AAW12841	pro-alpha2(1):(111
20	114.5	3.6	537	18 AAW12840	Mammalian Ena (Men
21	114.5	3.6	541	19 AAW37148	Human collagen (Ty
22	113.5	3.6	1838	15 AAR53257	Human ALP. Homo s
23	113	3.5	1274	20 AAW89253	Human retinoblasto
24	111.5	3.5	928	22 AAB28270	HNRCR protein sequ
25	111.5	3.5	2453	21 AAB12454	Murine WASP protei
26	111	3.5	520	22 AAB49336	BetaIII spectrin p
27	111	3.5	2390	20 AAY05494	Breast cancer asso
28	110	3.5	754	20 AAY07027	Human OREX ORF1999
29	110	3.5	2220	21 AAY42235	Full length calcin
30	110	3.5	2220	21 AAY54004	Rat Neural-Wiskott
31	109.5	3.4	501	19 AAW46890	Human OREX ORF1813
32	109	3.4	551	21 AAB42049	Human EXMAD-7 SEQ
33	109	3.4	795	22 AAB27229	Human retinoblasto
34	109	3.4	928	20 AAY01543	Human activated T-
35	109	3.4	1851	22 AAU00023	Human OREX ORF1983
36	108.5	3.4	782	21 AAB42219	Amino acid sequenc
37	108.5	3.4	2293	21 AAY69197	Human SLAP-130. H
38	108	3.4	783	20 AAW70586	Retinoblastoma sus
39	108	3.4	928	18 AAW09411	Infectious rubella
40	108	3.4	2205	16 AAR79048	Mouse IRS-2. Mus
41	107.5	3.4	1321	17 AAR96994	Retinoblastoma ppl
42	107	3.4	928	16 AAR71680	Recombinant ppl10R
43	107	3.4	928	16 AAR71681	Plasmid pRF1 rat l
44	106.5	3.3	678	8 AAP70380	Human cytoskeleton
45	106	3.3	324	21 AAY91948	

ALIGNMENTS

RESULT 1
ID AAW29749 standard; Protein; 610 AA.
XX
AC AAW29749;
XX
DT 11-JAN-1999 (first entry)
XX
DE Tumor suppressor protein menin.
XX
KW Familial multiple endocrine neoplasia type 1; FMEN1; MEN1;
KW menin; tumour suppressor gene; cancer; marker; diagnosis;
KW gene therapy; human.
XX
OS Homo sapiens.
XX
PN WO9839439-A1.
XX
PD 11-SEP-1998.
XX
PF 04-MAR-1998; 98WO-US04258.
XX
PR 05-MAR-1997; 97US-0040269.
XX
PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX
PI Agarwal SK, Burns AL, Chandrasekharappa SC, Collins FS;
PI Debelenko LV, Emmert-Buck MR, Guru SC, Liotta LA;
PI Lubensky IA, Manickam P, Marx SJ, Spiegel AM;
XX
XX WPI; 1998-506360/43.
DR
DR N-PSDB; AAV57415-16.
XX
PT Protein and gene associated with multiple endocrine neoplasia type 1

PT - useful in gene therapy and to diagnose sufferers of, and those
PT susceptible to, this condition by detecting protein absence or gene
PT mutation(s)

XX Claim 11; Page 56-57; 75pp; English.

CC This is the amino acid sequence of menin, a 67.5 kDa polypeptide
CC encoded by a novel human tumour suppressor gene, MEN1 (see AAV57416),
CC that is associated with multiple endocrine neoplasia type 1. Menin
CC exhibits no apparent amino acid sequence similarities to previously
CC known proteins. The lack of a functional menin polypeptide, either
CC by absence of the protein, its alteration and/or associated
CC mutations in the MEN1 gene, have been identified in individuals
CC suffering from multiple endocrine neoplasia type 1 (FEN1) and
CC identification of MEN1 provides a new window into the mechanism
CC of endocrine tumorigenesis, facilitates accurate early diagnosis of
CC MEN1 associated cancers, and provides preclinical identification of
CC individuals with the FEN1 syndrome, i.e. cancer-free individuals
CC that are at high risk of acquiring MEN1 associated tumours. It
CC also provides a basis for gene therapy. Menin can be expressed in
CC eukaryotic or prokaryotic host cells and used to identify menin
CC specific reagents (e.g. antibodies) useful for detecting menin in
CC humans cells or tissues, especially from patients suspected of
CC being at risk from multiple endocrine neoplasia type 1.

XX Sequence 610 AA;

Query Match 100.0%; Score 3185; DB 19; Length 610;
Best Local Similarity 100.0%; Pred. No. 3.5e-284;
Matches 610; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MGLKAQKTLFPLRSIDVYVRLFAELGREGPDVLVLSVLGFVEHFLAVNRVPTNVE 60
Db 1 mglkaqktlflrsidsdvrlfaaelgreepdlvllslvlgfvehflavnrviptnve 60
Qy 61 LTFQSPAPPPGGTLTFPPVADLSIIAALYARFTAQIRGAVDSLYPREGVSSRELVKK 120
Db 61 ltfqspapppggltfppvadlsiaalyarftaqirgavdlslypregvssrelvkk 120
Qy 121 VSDVWNSLSRSYFKRAHQISLFTGKFLDSSGVAFVAVGACQALGRDVLHALSED 180
Db 121 vsdvwnslsrsyfkdrahqlslftgkfldssgvafavvgacqalgrdvhlalsed 180
Qy 181 HAWVFGPNQSGTAEVTHWKGNDRRGQTVNAGVAERSWLYLKGSYMRCDRKMEVAFV 240
Db 181 hawvfgpngqtaevthwkgndrrgqtvnagvaerswlylkgysmrdrkmevafv 240
Qy 241 CAINPSIDLHFTDSLLELQLOKLLWLLYDLGLHLEPYPMALGNLADLELEPTGRPDPLT 300
Db 241 calnpsidlhtdslellqgkllwlllydlghlerypmalgnladleleptgrpdplt 300
Qy 301 LYHKGIASAKTYRDEHYIPYMYLAGYHCRNRVREALQAWADTATVTDYNYCREDEEI 360
Db 301 lyhkgiasaktyrdehyipymylagyhcnrnrvrealqawadtatviqdyncyredeei 360
Qy 361 YKEFEVANDVYIPNLLKEASLLLEAGEERPGESQSGTQSGSALQDPCEFAHLLRFYDGI 420
Db 361 ykefevandvipnllkeasllleageerpgesqsgtqsgsalqdpcecfahllrfydgil 420
Qy 421 CKWEGSPTVYLVHWGATFLVQSLGREGQVRQKRVIRVSREAAEAEPWGEAREGR 480
Db 421 ckweegsptvylvhwgatlflvqslgrfegqvrqkvrivsrtaeaaeapwgearegr 480
Qy 481 RGPRESKPEPPPPKPKALDKGLTGCGAVSGPPRPGTVAGTARGCGGSTAQVPAP 540
Db 481 rgpreskpeppppkpkaldkgltgqgavsgpprpgtvgatargcggstaqvpap 540
Qy 541 AASPPPEPVLTFQSEKMGKMKELLVATKINSSAIKILQTAOSQVQMKKQKVSFSDYTL 600
Db 541 aaspppepvltfqsekmgkmkellvatkinssaiikilqtaqsqvqmkkqkvspsdytl 600

Qy 601 SFLKRRQRKGL 610
Db 601 sflkrqrkgll 610
RESULT 2
AAW86348
ID AAW86348 standard; Protein; 566 AA.
XX AC AAW86348;
XX 15-MAR-1999 (first entry)
XX Human multiple endocrine receptor neoplasia type 1 protein.
XX Human; multiple endocrine neoplasia type 1 protein; MEN1-1;
XX tumorigenesis; cancer; multiple endocrine neoplasia; gene therapy;
XX detection; diagnosis; drug screening.
XX Homo sapiens.
XX Key Location/Qualifiers
XX Misc-difference 137 /label= unknown
XX Misc-difference 161 /note= "encoded by YGG"
XX Misc-difference 274 /label= unknown
XX Misc-difference 274 /note= "encoded by NNT"
XX Misc-difference 312 /label= unknown
XX Misc-difference 312 /note= "encoded by RGG"
XX Misc-difference 417 /label= unknown
XX Misc-difference 417 /note= "encoded by SCC"
XX Misc-difference 421 /label= unknown
XX Misc-difference 421 /note= "encoded by TNG"
XX Misc-difference 424 /label= unknown
XX Misc-difference 424 /note= "encoded by NGC"
XX Misc-difference 428 /label= unknown
XX Misc-difference 428 /note= "encoded by GAN"
XX Misc-difference 435 /label= unknown
XX Misc-difference 435 /note= "encoded by CNG"
XX Misc-difference 443 /label= unknown
XX Misc-difference 443 /note= "encoded by ANC"
XX W09854324-A1.
XX 03-DEC-1998.
XX 29-MAY-1998; 98WO-US10957.
XX 29-MAY-1997; 97US-0865337.
XX (INCY-) INCYTE PHARM INC.
XX Au-Young J, Covitz PA, Murry LE, Tang YT;
XX WPI; 1999-059839/05.
XX N-PSDB; AAV80659.
XX New isolated multiple endocrine neoplasia type 1 protein - used to
XX develop products for the diagnosis, treatment and prevention of
XX cancers and multiple endocrine neoplasia.
XX Claim 1; Fig 1; 67pp; English.
PS

XX The present sequence is human multiple endocrine neoplasia type 1
CC protein (MEND-1). MEND-1 plays a role in multiple endocrine neoplasia
CC when one or both normal genetic copies of MEND-1 are mutated and no
CC longer able to suppress tumorigenesis. MEND-1 can be used for treating
CC or preventing cancers and multiple endocrine neoplasia. MEND-1
CC polynucleotides can also be used for gene therapy. Products from the
CC present invention can also be used for detection, diagnosis and drug
CC screening.
XX Sequence 566 AA;
SQ
Query Match 87.2%; Score 2778; DB 20; Length 566;
Best Local Similarity 89.5%; Pred. No. 8.7e-247;
Matches 547; Conservative 2; Mismatches 16; Indels 46; Gaps 4;
QY 1 MGLKAQAQKTLFPLRSDDVVRFAAELGREEDPLVLLSLVLGFEVHFLAVNRVIPNVPE 60
Db 1 mglkaaqktlflrsiddvvrfaaelgreepdlvllslvlgfvehflavnrviptnvpe 60
QY 61 LTFQSPADPPGGLNFPVADLSIIAALYARETAQIRGAVDLSLYPREGVSSRELVKK 120
Db 61 ltfqspapppggltyfvpvadisliiaalyarftaqirgavdlslypreggvssrelvkk 120
QY 121 VSDVINSLRSYFKDRAHIQSLFSFITGKLDSSGVAFVAVGACQALGLRDVHLASED 180
Db 121 vsdvinslrsyfkdxahiqslfsfitgkld-sgvafa----- 159
QY 181 HAWVFGNGEQTAETVHWGKGNEDRRGQTVNAGVAERSWLYLKGYMCRDKMEVAFMV 240
Db 160 -----vxgpngeqtaetvhwgkgnedrrgqtvnagvaerswlylkgymsrcdkmevafmv 215
QY 241 CATNPSTDLHTDSLELQLOOKLLWLLYDLGHLERYPMALGNLADLEELEPTGCRDPLT 300
Db 216 catnpsidldhtdslelqgkllwlllydlghlerypmaignladileetptgrrpdpaxa 275
QY 301 LYHKGIASAKTYRDEHIYPYMYLAGYHCRNRNVREALQAWA-DTATVIQDYNCREDEE 359
Db 276 llhkgiasaktyrdehiypymylagyhcrnrnvrealqalgrdtatliqdyncredee 335
QY 360 IYKEFEVANDVPNLLKEAASLLERAGEPRGOSQGTQSQGALQDPECFALLRFPYDG 419
Db 336 iykefevandvipnllkeaslleageerpgesqgtqsgsalqdpcecfahllrfydg 395
QY 420 ICKWEGSPPTVLHVGWATFLVQSLGRFEGQVQKRVIVSREAEAEAEPEWGEAREGR 479
Db 396 ickweegspcpvlhvqwapflxgslrxfqgqvqkvrivxrdaea----- 441
QY 480 RRGPRRESKPEEPPPKPALDKLGTGQGVAGSGPPRKPPTVAGTARGPEGGSTAQVPA 539
Db 442 -----exkpeepppkpaldklgtgqgvagsgpprkpvtvagtargpeggstaqvpa 495
QY 540 PAASPPPEGVPLTFQSEKMGKMKELLVATKINSALKQLTATQSOVMKKQKYSTPDSYT 599
Db 496 paaspppegvpltfqsekmgmellvatkinsalklqtatqsgvmqkqkystpsdyt 555
QY 600 LSLFKRQRKGL 610
Db 556 lsflkrqakgl 566
RESULT 3
ID AAW77287
XX AAW77287 standard; Protein; 1151 AA.
XX
AC AAW77287;
XX
DT 20-NOV-1998 (first entry)
XX
DE Zebrafish differentiation enhancing factor 1 protein.
XX
KW Zebrafish; differentiation enhancing factor; ankyrin repeat; C2 domain;

KW SH3 consensus binding sequence; pleckstrin homology domain; adipogenesis;
KW neurogenesis; hyperplastic disease; neoplastic disease; nervous system.
XX Brachydanio rerio.
XX
XX
PN W09836065-A1.
XX
PD 20-AUG-1998.
XX
XX
PF 13-FEB-1998; 98WO-US02724.
XX
PR 14-FEB-1997; 97US-0038191.
XX
PA (DAND) DANA FARBER CANCER INST INC.
XX
PI Chan J, Harris DF, Hu E, King FJ, Spiegelman B;
PI Thomas RM;
XX
XX WPI; 1998-467173/40.
DR N-PSDB; AAV59104.
XX
XX New nucleic acid encoding differentiation enhancing factor - used
PT particularly to regulate adipogenesis and neurogenesis, e.g. for
PT treating tumours and neurological disease
XX
PS Claim 3; Fig 12; 203pp; English.
XX
XX The differentiation enhancing factors (DEF), comprise at least one each
CC of SH3 consensus binding sequence, ankyrin repeat, pleckstrin homology
CC domain and C2 domain. The DEF induce adipogenesis or neurogenesis; they
CC are mediators of SH3-domain dependent signalling and may be involved in
CC cellular gene expression, cytoskeletal architecture, protein trafficking,
CC endocytosis or adhesion, migration, proliferation and differentiation of
CC cells. Typical applications of DEF and agents that modulate
CC interaction between the protein and its ligand, or of nucleic acid
CC expressing them, are treatment of hyperplastic and neoplastic diseases
CC (a wide range of solid tumours and leukaemias), including metastases; for
CC in vitro induction of differentiation of neural crest cells to neurons,
CC glial cells etc.; for increasing neuron survival, and inducing cell
CC repair, in the nervous system (e.g. treatment of traumatic injury,
CC stroke, Alzheimer's, Parkinson's or Huntington's diseases, amyotrophic
CC lateral sclerosis, multiple sclerosis etc.).
XX Sequence 1151 AA;
Query Match 4.1%; Score 131; DB 19; Length 1151;
Best Local Similarity 21.4%; Pred. No. 0.0097;
Matches 118; Conservative 73; Mismatches 220; Indels 140; Gaps 26;
QY 119 KKYSDVINWNSLSRYF---KDRAHIQSLFSFITGKLDSSGVAFVAVGACQALGLRDVHL 175
Db 382 kkcfdli--shnrythfqaedeqefviwsvltnskealnmafrgeqsagdsledltk 439
QY 176 ALSEDHAWVFGNGEQTAETVHWGKGNEDRRGQTVNAGVAERSWLYLKGYMCRDKME 235
Db 440 aliiedvlri-----pgne-----vccdcgvpepkwlstnlgilc----- 474
QY 236 VAFWVCA-INPSTDLH---TDSLELQLOOKLLWLLYDLGHLERYPMALGNLADLEELEP 291
Db 475 ---iecsghremgvhisriqsmeldkigtseillaknvgnssfnellegnl-----p 524
QY 292 TPG-RPDPDLTYHKGIASAKTYRDE---EHYIPYMYLAGYHCRNRNVREALQAWADTATV 347
Db 525 spspkpa-----ssdmterkeyinakyvehrfarrtattatargdlyeavrt-rdlmal 579
QY 348 IQDYNCREDEEITYKEFFEIVANDVIPNLL-----KEAASLLEAGEERKPEGSQGTQS 399
Db 580 iqly---adgvelmfppeagqdpgetalhfavrtsqtslhldvflvqnsqgtldrqtcs 636
QY 400 QGSAL-----QDPECFAHLLRFYDGICKWEESPTP-----VLHVGWATFLVQ-SLGR 446
Db 637 gnaalhyccyckpeciklllrgkpsidlvnqngetaldiarlrnvqceallveaagr 696

XX PS Claim 1; Page 17-21; 28pp; English.

XX CC This invention relates to an immediate early herpes-simplex-virus type 2

CC (HSV-2) infected cell protein 4 (ICP4) recognised by human cytotoxic T

CC cells. HSV-2 ICP4 protein is recognized by cytotoxic T-lymphocyte (CTL)

CC cells in humans and is used in vaccines for therapeutically or

CC prophylactically treating HSV infections. Pharmaceutical compositions of

CC HSV-2 ICP4 protein may be used to treat patients suffering from HSV

CC infections, to prevent or decrease recurrent herpes disease, frequency,

CC severity and duration of episodes. The present sequence represents the

XX HSV-2 ICP4 amino acid sequence of the invention.

SQ Sequence 1318 AA;

Query Match 3.8%; Score 120.5; DB 22; Length 1318;

Best Local Similarity 25.4%; Pred. No. 0.11;

Matches 62; Conservative 19; Mismatches 108; Indels 55; Gaps 9;

QY 354 CREDEIYKEFEVANDVIPNLLKEAASLLEAGERPGQSGQSALQDPECF AHL 413

DB 644 crgilealaegfdgdaavpgl----agarpaaprrpgagaaphada---prlrawl 596

QY 414 --LRF-----YDICKWEEGSPPTVLHVGMATFLVOSLG----- 445

DB 697 relfrvdalvmlrldrvaggseaaavaavrsavslvagalgpalprsprllssaaaaa 756

QY 446 ---REGQVRQKV---RIVSRGAEEAEEPEWGEAREGRRGPRRESKPEPPPKK-- 497

DB 757 adllfqngslrplladtvaadslaapasap--rearkrspaparapggaprrpkksr 814

QY 498 ----PALDKGLGTGGGAVSGPPRPPGVAGTAR-----GPEGGSTAOVPAPASPPE 547

DB 815 adaprraaappagaappaptpprrprpaaltrpaegdpqggwrrqpgpshtcps 874

QY 548 GPVL 551

DB 875 aaal 878

RESULT 6

ID AAB29662 standard; Protein; 1583 AA.

AC AAB29662;

XX DT 23-FEB-2001 (first entry)

XX DE Human tyrosine phosphatase HD-PTP fragment.

XX KW Human; histidine domain-protein tyrosine phosphatase; HD-PTP;

XX KW chromosome 3p21.3; gene deletion; tumour suppressor; cytostatic;

XX KW lung cancer; tumour; gene therapy; diagnosis; recombinant production;

XX KW anticancer.

XX OS Homo sapiens.

XX PN WO200063392-A1.

XX PD 26-OCT-2000.

XX PF 14-APR-2000; 2000WO-JP02455.

XX PR 16-APR-1999; 99JP-0108842.

XX PA (KYOW) KYOWA HAKKO KOGYO KK.

XX PI Shimizu K;

XX DR WPI; 2000-672740/65.

XX DR N-PSDB; AAC81225.

PT Human tyrosine phosphatase with oncostatic activity encoded by a gene

PT frequently deleted in lung cancer, useful for treatment and diagnosis

PT of tumors

XX Example 2; Page 82-98; 134pp; Japanese.

XX The invention relates to a novel human tyrosine phosphatase, histidine

CC domain-protein tyrosine phosphatase (HD-PTP; AAB29661) and to human

CC HD-PTP nucleic acids (AAC81224, AAC81225, AAC81262, AAC81263). The

CC HD-PTP gene is located on chromosome 3p21.3. This region is frequently

CC found to be deleted in lung cancers, and is therefore thought to contain

CC a tumour suppressor gene. The invention also relates to expression

CC vectors and host cells containing human HD-PTP nucleic acids; the

CC recombinant production of HD-PTP; anticancer drugs containing HD-PTP;

CC gene therapy compositions containing DNA encoding HD-PTP; diagnostic

CC reagents containing HD-PTP oligonucleotides; antibodies specific for

CC HD-PTP; and an immunoassay method using HD-PTP-specific antibodies for

CC use in cancer diagnosis and investigation. HD-PTP proteins, nucleic acids

CC and antibodies may be used in the treatment, investigation and diagnosis

CC of cancers, particularly those of the lung. The present sequence

XX represents a substantial portion of the human HD-PTP protein.

SQ Sequence 1583 AA;

Query Match 3.7%; Score 118.5; DB 21; Length 1583;

Best Local Similarity 21.8%; Pred. No. 0.22;

Matches 138; Conservative 73; Mismatches 224; Indels 197; Gaps 34;

QY 24 AAELGREPDVLLSL-----VLG-----FVEHFLAVNRVPTNVP EL-TFQP- 65

DB 230 aiklakgqddtvgdalrftmdvlggkynsakknddfyh-----eavpalclqpv 280

QY 66 --SPAPDP--PGGLTFPPVADLSIIAALYARFQAIRGAVDLSLYPREGGVSSRELKVKYS 122

DB 281 kgaplrvkplpvnpdpavtvgpddifaklvpmahaas-----slyseekaklremmakie 335

QY 123 DVIWNSLSRSYFKDRAHIQSLFSFITGTRKLDSSGV-----AFA-----VVGACOA L- 168

DB 336 d-----knevldqfmsmqldpbtvndldayshippqlmekcaalsvrpd 380

QY 169 -----GLRDVHLALSEDHAWVFGPNCEQTAETVTHGKGNEDRR 207

DB 381 tvrnlvqsmqvlgvftdveaslkdrilleed-----elleqkfgevvgqagai 430

QY 208 GQTVNAGVAE--RSWLYLKGSYMRCDRKMVAFWCAINPSIDLHTDLSLELLOLQKLLW 265

DB 431 sitskaelaevrrew----akymevhek--asftnselhramnlhvgnrlrls----- 477

QY 266 LLYDLGHLERYPMALGNLADLELEPTP--GRDPDLTYH-KGIASAKTYVYRDEHIVPYM 322

DB 478 -----gpldqvrval-----ptpalspedkavlgqlkrilakvqemdrgrv---- 518

QY 323 YLAGYHCRNRNVREALQAWADATVI-QDYNVYCRE-DEEYIYKEFEFVANDVIPNLLKE-- 378

DB 519 -----sleqqlireliqkdditaslvttdhsemkklifeeqkkydqklyleqnlqaadr 572

QY 379 -AASLLEAGEERPEEQ---SOGTOSQSALQ---DPECFAHLLREYDGTCKWEESPTP 430

DB 573 vicalteanvgaavrvrvisldldqkwnstlqtlvasyayedlmk-----ksqegrdfy 626

QY 431 V-LHVGWATFLVQSLGRFEGQVRQKVRIVSR-----EAEAAEAE 468

DB 627 adleskvaallertqstcqareaaqqldrelkkpprptapklprreeseaveag 686

QY 469 EPWGEAEARE---GRRRGPR-----RESKPEEPPPKPKPALDKGLGTGGAVSGP RPKPP 519

DB 687 dp-peeelrsllpdmvagrplpdtflgsatplhfpssfpfs---stgpgplylsqp--lpp 740

QY 520 GTVAGTARGPEGGSTAQVPAPASPPEGPVL 551

DB 741 gtyS-----gptqliprapgppampvapgal 768

AC AAW50192;
XX 28-JUL-1998 (first entry)
XX Amino acid sequence of salivary protein CON-1.
DE Salivary glycoprotein; CON-1; CON-2; alpha-glucosidase inhibitor;
KW retrovirus; inhibition; diabetes; HIV.
XX Homo sapiens.
OS
XX WO9809981-A1.
PN 12-MAR-1998.
XX 08-SEP-1997; 97WO-US15799.
PF 09-SEP-1996; 96US-0024712.
PR (WISC) WISCONSIN ALUMNI RES FOUND.
XX Azen EA, Pan D;
XX WPI; 1998-193547/17.
DR N-PSDB; AAV22062.
XX Isolated salivary glyco-protein CON-1 and CON-2 compositions - which
XX have alpha-glucosidase inhibitory activity, useful for treating
PT diabetes or retrovirus, particularly HIV infection
PT
XX Disclosure; Page 32; 54pp; English.
PS
XX This amino acid sequence is of the salivary glycoprotein CON-1,
CC apotent alpha-glucosidase inhibitor, which is useful in preventing
CC cellular penetration of retroviruses. In the method of the invention
CC CON-1 inhibits alpha-glucosidase processing of the retroviruses
CC required for proper engagement of the virion with its cellular
CC receptor. The salivary glycoproteins CON-1 and CON-2 and derivatives,
CC have AGS inhibitory activity and can be used to treat patients with
CC diabetes or patients infected with retroviruses such as HIV.
CC
XX Sequence 124 AA;
SQ
Query Match 3.7%; Score 117.5; DB 19; Length 124;
Best Local Similarity 32.3%; Pred. No. 0.0053;
Matches 31; Conservative 9; Mismatches, 25; Indels 31; Gaps 5;
Qy 469 EPWGEAREGRRRRPRRE--SKPEPPPPKPKALDKLGTGGQAVSGPP-----R 516
Db 16 qpqgppppgkpggppggnkpggpppgk-----qgpppggdnksqsr 62
Qy 517 KPGCTVAGTARGEGGTAQVPAP----AASPPPEG 548
Db 63 sppgkpgppp--pdggnqpggpppppgkpggpppgq 96
RESULT 11
AAY06309
ID AAY06309 standard; Protein; 1566 AA.
XX
AC AAY06309;
XX
XX 06-SEP-1999 (first entry)
XX Human p53 regulatory protein RB18A.
XX RB18A; p53 regulatory protein; apoptosis; neoplasia; inflammation;
KW wound healing; graft rejection; reperfusion injury;
KW myocardial infarction; stroke; traumatic brain injury;
KW neurodegenerative disease; ischaemia; toxemia; infection; AIDS;
KW hepatitis; breast cancer; ovarian cancer; colon cancer; diagnosis;
KW therapy; human.

XX Homo sapiens.
OS
XX Key
FH Domain
FT Location/Qualifiers
FT /note= "DNA binding domain; this polypeptide is
FT specifically claimed in Claim 4"
FT 1234..1406
FT /note= "p53 binding and homo-oligomerisation
FT domains; this polypeptide is specifically
FT claimed in Claim 4"
FT 955..958
FT /note= "nuclear localisation signal"
FT 979..982
FT /note= "nuclear localisation signal"
FT 996..999
FT /note= "nuclear localisation signal"
FT 1483..1493
FT /note= "nuclear localisation signal" ;
FT 436..1566
FT /note= "this polypeptide is specifically claimed
FT in Claim 4"
FT 436..1228
FT /note= "this polypeptide is specifically claimed
FT in Claim 4"
FT 436..927
FT /note= "this polypeptide is specifically claimed
FT in Claim 4"
FT 1537..1566
FT /note= "this polypeptide is specifically claimed
FT in Claim 4"
FT 1234..1566
FT /note= "this polypeptide is specifically claimed
FT in Claim 4"
XX WO9931231-A1.
XX 24-JUN-1999.
XX 14-DEC-1998; 98WO-EP08560.
XX 15-DEC-1997; 97EP-0403051.
XX (INRM) INSERM INST NAT SANTE & RECH MEDICALE.
XX Frade R;
XX WPI; 1999-3951177/33.
DR N-PSDB; AAX59124.
XX New p53 regulatory protein (RB18A) useful as, e.g. sources of probes
XX and primers to detect the transcription rate and abundance of RB18A
XX mRNA in lymphocytes
XX Claim 5; Fig 3; 87pp; English.
XX This sequence represents human RB18A (recognized by PAb1801
CC antibody), a novel 205 kDa p53 regulatory protein that inhibits
CC p53-induced apoptosis. RB18A cDNA (see AAX59124) was isolated from
CC a human heart cDNA library. RB18A shares some antigenic epitopes
CC with p53. It regulates the sequence-specific DNA binding function
CC of p53. Its C-terminal domain induces a dose-dependent
CC reassociation of DNA double strands, stimulates in vitro the
CC sequence-specific DNA binding of p53, and functionally interacts
CC and stimulates in vivo the transactivation activity of p53. RB18A
CC polypeptides can be used to block the G1 phase of the cell cycle
CC and/or regulate apoptosis, and so control cell growth. This makes
CC them good candidates for antineoplastic therapy. Pharmaceutical
CC compositions comprising RB18A polypeptides or polynucleotides are
CC useful for preventing or treating a variety of human and veterinary
CC diseases, e.g. neoplasia, inflammation, wound healing, graft
CC rejection, reperfusion injury, myocardial infarction, stroke,
CC traumatic brain injury, neurodegenerative diseases, ageing,

CC ischaemia, toxemia, infection, AIDS and hepatitis. Preferably, they are used as antineoplastic compositions, or compositions directed against any other cell proliferation disease, especially in the treatment of colon, breast or ovarian cancer.

XX Sequence 1566 AA;

Query Match 3.7%; Score 117.5; DB 20; Length 1566;
Best Local Similarity 22.4%; Pred. No. 0.27;
Matches 72; Conservative 33; Mismatches 128; Indels 89; Gaps 13;

QY 341 WADTATVIQDYNVCREDEIYKEFEFVANDVIPNLKEAASLLEAGEERPEGEOSQ----- 395
Db 820 ytdpadliadaagsspsdptnhfhdvdfnplll-nsqsgsfgeeyfdesqsgdnd 878

QY 396 ---GTOSQ-----GSALQDPECFALLRRF 416
Db 879 dfkgfasqalntlgvmlgdngetkfgnngadtvdfsliisvagkalapadimeh---- 934

QY 417 YDGICKWEGSPPTVLHVGMWATFLVQSLGRFEGQVRQKRVISREAEAEPEWGEAR 476
Db 935 -----hsgsgpllttg-----dlgkektqkrvk---egngtsnstlsgp-glask 976

QY 477 EGRR-RGPRRESKPEPPPPKPPALDKGLGTGGGAVSGPPRPKPGTVAGTARGPEGSTA 535
Db 977 pgkrstrpsndgskdkppkrkkadtekgspshss-snrpftpp-tstgskspgsagrs 1034

QY 536 QVPAPASPP-----PEGPVLTFQSEKMGKMKELLVATKINSSAIK-----LQL 579
Db 1035 qtpgvgatppipkitiqikgtvmv---gkpsghsqtssgsvssgskshshssss 1091

QY 580 TAQSQVMKKQKVSTPSTDTLS 601
Db 1092 sastsgkmkssksegssskls 1113

RESULT 12
AAY69669
ID AAY69669 standard; Protein; 1581 AA.
XX AC AAY69669;
XX DT 08-MAY-2000 (first entry)
XX DE Human thyroid receptor-associated protein TRAP220.
XX KW Thyroid receptor-associated protein; TRAP220; TRAP complex; coactivator;
KW nuclear hormone receptor; thyroid receptor; vitamin D receptor;
KW oestrogen receptor; mineralcorticoid receptor;
KW peroxisome proliferation-activated receptor; LXXLL motif; drug screening;
XX detection.

OS Homo sapiens.
XX FH Key Location/Qualifiers
FT Region 604..608
FT /note= "LXXLL motif 1"
FT Region 645..649
FT /note= "LXXLL motif 2"

XX W0200001820-A2.
XX PN 13-JAN-2000.
XX PD 01-JUL-1999; 99WO-US15052.
XX PF 06-JUL-1998; 98US-0110517.
XX PR (UYRQ) UNIV ROCKEFELLER.
XX PA Roeder RG, Fondell JD, Xingyuan C, Ito M;
XX PI

DR WPI: 2000-147418/13.
DR N-PSDB; AA287101.
XX
PT New isolated Thyroid Receptor-Associated Proteins which act as nuclear
PT hormone receptor coactivators used for identifying modulators of
PT hormones or nuclear hormone receptors
XX
PS Claim 9; Fig 5A; 114pp; English.
XX
CC The present sequence represents human thyroid receptor-associated
CC protein TRAP220. TRAP220 is a member of a complex of TRAPs which
CC act as coactivators for nuclear hormone receptors, binding
CC to such receptors in a ligand-dependent manner and are required for
CC functional interactions between the receptor and genes whose
CC transcription is regulated by these receptors. Nuclear hormone receptors
CC include thyroid receptors (TRs), vitamin D receptors (VDRs), oestrogen
CC receptors (ERs), mineralcorticoid receptors (MRs) and peroxisome
CC proliferation-activated receptors (PPARs). TRAP220 contains two of the
CC LXXLL motifs that have been implicated in nuclear hormone receptor-
CC coactivator interactions. TRAP220, and a second protein of the
CC invention, TRAP100 (AAY69670), and nucleotides encoding these proteins
CC may be used to modulate the activity of a nuclear hormone receptor, or to
CC screen for agents that modulate receptor or hormone activity. Proteins,
CC nucleic acids and antibodies may also be used therapeutically and for
CC detection of TRAP220 and TRAP100 or their associated nucleotides.
XX
SQ Sequence 1581 AA;

Query Match 3.7%; Score 117.5; DB 21; Length 1581;
Best Local Similarity 22.4%; Pred. No. 0.28;
Matches 72; Conservative 33; Mismatches 128; Indels 89; Gaps 13;

QY 341 WADTATVIQDYNVCREDEIYKEFEFVANDVIPNLKEAASLLEAGEERPEGEOSQ----- 395
Db 835 ytdpadliadaagsspsdptnhfhdvdfnplll-nsqsgsfgeeyfdesqsgdnd 893

QY 396 ---GTOSQ-----GSALQDPECFALLRRF 416
Db 894 dfkgfasqalntlgvmlgdngetkfgnngadtvdfsliisvagkalapadimeh---- 949

QY 417 YDGICKWEGSPPTVLHVGMWATFLVQSLGRFEGQVRQKRVISREAEAEPEWGEAR 476
Db 950 -----hsgsgpllttg-----dlgkektqkrvk---egngtsnstlsgp-glask 991

QY 477 EGRR-RGPRRESKPEPPPPKPPALDKGLGTGGGAVSGPPRPKPGTVAGTARGPEGSTA 535
Db 992 pgkrstrpsndgskdkppkrkkadtekgspshss-snrpftpp-tstgskspgsagrs 1049

QY 536 QVPAPASPP-----PEGPVLTFQSEKMGKMKELLVATKINSSAIK-----LQL 579
Db 1050 qtpgvgatppipkitiqikgtvmv---gkpsghsqtssgsvssgskshshssss 1106

QY 580 TAQSQVMKKQKVSTPSTDTLS 601
Db 1107 sastsgkmkssksegssskls 1128

RESULT 13
AAY69671
ID AAY69671 standard; Protein; 1581 AA.
XX AC AAY69671;
XX DT 08-MAY-2000 (first entry)
XX DE Human thyroid receptor-associated protein TRAP220 mutant a.
XX KW Thyroid receptor-associated protein; TRAP220; TRAP100; coactivator;
KW TRAP complex; nuclear hormone receptor; thyroid receptor;
KW vitamin D receptor; oestrogen receptor; mineralcorticoid receptor;
KW peroxisome proliferation-activated receptor; LXXLL motif; drug screening;
KW detection; mutant; mutein.

XX Homo sapiens.
OS Synthetic.
FH Key Location/Qualifiers
FT Region 604..608
FT /note= "Mutated LXXLL motif (LXXAA)"
FT 645..649
FT /note= "LXXLL motif 2"
XX
XX WO200001820-A2.
XX
XX 13-JAN-2000.
XX
XX 01-JUL-1999; 99WO-USI5052.
XX
XX 06-JUL-1998; 98US-0110517.
XX
XX (UYRQ) UNIV ROCKEFELLER.
XX
XX Roeder RG, Fondell JD, Xingyuan C, Ito M;
XX
XX WPI: 2000-147418/13.
XX
XX New isolated Thyroid Receptor-Associated Proteins which act as nuclear
XX hormone receptor coactivators used for identifying modulators of
XX hormones or nuclear hormone receptors
XX
XX Example; Page -: 114pp; English.
XX
XX The invention relates to human thyroid receptor-associated proteins
XX TRAP220 (AAV69669) and TRAP100 (AAV69670) and nucleotides encoding them
XX (AAZ87101-287102). TRAP220 and TRAP100 are members of a complex of TRAPS
XX which act as coactivators for nuclear hormone receptors, binding
XX to such receptors in a ligand-dependent manner and are required for
XX functional interactions between the receptor and genes whose
XX transcription is regulated by these receptors. Nuclear hormone receptors
XX include thyroid receptors (TRs), vitamin D receptors (VDRs), oestrogen
XX receptors (ERs), mineralocorticoid receptors (MRs) and peroxisome
XX proliferator-activated receptors (PPARs). TRAP220 contains two of the
XX LXXLL motifs that have been implicated in nuclear hormone receptor-
XX coactivator interactions, while TRAP100 contains six of these motifs.
XX TRAP220 and TRAP100, and their associated nucleotides, may be used to
XX modulate the activity of a nuclear hormone receptor, or to screen for
XX agents that modulate receptor or hormone activity. Proteins, nucleic
XX acids and antibodies may also be used therapeutically and for detection
XX of TRAP220 and TRAP100 or their associated nucleotides. Sequences
XX of TRAP220 and TRAP100 represent TRAP220 mutants generated in an
XX exemplification of the present invention to test the functional role of
XX TRAP220 and the contribution of the LXXLL motifs.
XX Note: The present sequence is not shown in the specification, but is
XX derived from the wild-type TRAP220 protein sequence given in fig 5A.
XX
SQ Sequence 1581 AA;

Query Match 3.7%; Score 117.5; DB 21; Length 1581;
Best Local Similarity 22.4%; Pred. No. 0.28;
Matches 72; Conservative 33; Mismatches 128; Indels 89; Gaps 13;

Qy 341 WADTATVTDYNYCREDEIYKEFEVANDVIPNLLKEAALLRGERPGEQSQ----- 395
Db 835 ytdpadliadaagspsdsptnhffdgvdnpdl1-nsqsgsgfgeeyfdesqsgdnd 893
Qy 396 ---GFSQS-----GSALQDPECFAHLLRF 416
Db 894 dfkgfasqalntlgyplmgdnggetckfgnngadtvdfsiisvagkalapadlmeh---- 949
Qy 417 YDGICKWKEGSPTPVLHVHGWATFLVQSLGRFEGQVQRKRVIVSREAAEAEEEPWGEAR 476
Db 950 -----hsqsggpllttg-----dlgkektqkrvk---egnqtsnstlsgp-gldsk 991
Qy 477 EGRR-RGPRRESKPEEPPPPKKPALDKGLGTGGQAVSGPGRKPPGTAGTARGPEGGSTA 535

Db 992 pgrsrtpsndgskdkppkrkkadteggkspshs-snrptpp-tstggskspgsgrs 1049
Qy 536 QVPAPASPP-----BEGPVLTFQSEKMKMKELLVATKINSSAIK-----LQL 579
Db 1050 qtpgvatpplpkltlqipkgvmv---gkpsshsqtstsgvsssgskshshsssss 1106
Qy 580 TAOQVQMKKOKVSTPSPDYTLS 601
Db 1107 satsgmkksksegssskls 1128
RESULT 14
AAV69672
ID AAV69672 standard; Protein; 1581 AA.
XX
XX AAV69672;
XX
XX 08-MAY-2000 (first entry)
XX
XX Human thyroid receptor-associated protein TRAP220 mutant b.
XX
XX Thyroid receptor-associated protein; TRAP220; TRAP100; coactivator;
XX TRAP complex; nuclear hormone receptor; thyroid receptor;
XX vitamin D receptor; oestrogen receptor; mineralocorticoid receptor;
XX peroxisome proliferation-activated receptor; LXXLL motif; drug screening;
XX detection; mutant; mutein.
XX
XX Homo sapiens.
OS Synthetic.
XX
XX Key Location/Qualifiers
XX Region 604..608
XX /note= "LXXLL motif 1"
XX Region 645..649
XX /note= "Mutated LXXLL motif (LXXAA)"
XX
XX WO200001820-A2.
XX
XX 13-JAN-2000.
XX
XX 01-JUL-1999; 99WO-USI5052.
XX
XX 06-JUL-1998; 98US-0110517.
XX
XX (UYRQ) UNIV ROCKEFELLER.
XX
XX Roeder RG, Fondell JD, Xingyuan C, Ito M;
XX
XX WPI: 2000-147418/13.
XX
XX New isolated Thyroid Receptor-Associated Proteins which act as nuclear
XX hormone receptor coactivators used for identifying modulators of
XX hormones or nuclear hormone receptors
XX
XX Example; Page -: 114pp; English.
XX
XX The invention relates to human thyroid receptor-associated proteins
XX TRAP220 (AAV69669) and TRAP100 (AAV69670) and nucleotides encoding them
XX (AAZ87101-287102). TRAP220 and TRAP100 are members of a complex of TRAPS
XX which act as coactivators for nuclear hormone receptors, binding
XX to such receptors in a ligand-dependent manner and are required for
XX functional interactions between the receptor and genes whose
XX transcription is regulated by these receptors. Nuclear hormone receptors
XX include thyroid receptors (TRs), vitamin D receptors (VDRs), oestrogen
XX receptors (ERs), mineralocorticoid receptors (MRs) and peroxisome
XX proliferator-activated receptors (PPARs). TRAP220 contains two of the
XX LXXLL motifs that have been implicated in nuclear hormone receptor-
XX coactivator interactions, while TRAP100 contains six of these motifs.
XX TRAP220 and TRAP100, and their associated nucleotides, may be used to
XX modulate the activity of a nuclear hormone receptor, or to screen for
XX agents that modulate receptor or hormone activity. Proteins, nucleic
XX acids and antibodies may also be used therapeutically and for detection

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 03:45:17 ; Search time 8321.72 Seconds
(without alignments)
10427.796 Million cell updates/sec

Title: US-09-380-337-3
Perfect score: 9180
Sequence: 1 CTGGTCVTGAACCTCTGGCC.....AGCCACGGGCGCGCGCCCG 9180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*	
1: gb_est1:*	em_esthum10:*
2: gb_est2:*	em_esthum11:*
3: gb_est3:*	em_esthum12:*
4: gb_est4:*	em_esthum13:*
5: gb_est5:*	em_esthum14:*
6: gb_est6:*	em_esthum15:*
7: gb_est7:*	em_esthum16:*
8: gb_est8:*	em_esthum17:*
9: gb_est9:*	em_esthum18:*
10: gb_est10:*	em_esthum19:*
11: gb_est11:*	em_esthum20:*
12: gb_est12:*	em_esthum21:*
13: gb_est13:*	em_esthum22:*
14: gb_est14:*	em_esthum23:*
15: gb_est15:*	em_esthum24:*
16: gb_est16:*	em_esthum25:*
17: gb_est17:*	em_esthum26:*
18: gb_est18:*	em_esthum27:*
19: gb_est19:*	em_esthum28:*
20: gb_est20:*	em_estin1:*
21: gb_est21:*	em_estin2:*
22: gb_est22:*	em_estin3:*
23: gb_est23:*	em_estin4:*
24: gb_est24:*	em_estin5:*
25: gb_est33:*	em_estom1:*
26: gb_est34:*	em_estom2:*
27: gb_est35:*	em_estov1:*
28: gb_est36:*	em_estov2:*
29: gb_est37:*	em_estp11:*
30: gb_est38:*	em_estp12:*
31: gb_est39:*	em_estp13:*
32: gb_est40:*	em_estp14:*
33: em_estba:*	em_estp15:*
34: em_estfun:*	em_estp16:*
35: em_esthum1:*	em_estp17:*
36: em_esthum2:*	em_estp18:*
37: em_esthum3:*	em_estp19:*
38: em_esthum4:*	em_estp110:*
39: em_esthum5:*	em_estro1:*
40: em_esthum6:*	em_estro2:*
41: em_esthum7:*	em_estro3:*
42: em_esthum8:*	em_estro4:*
43: em_esthum9:*	em_estro5:*
	em_estro6:*
	em_estro7:*
	em_estro8:*
	em_estro9:*
	em_estro10:*
	em_estro11:*
	em_estro12:*
	em_estro13:*
	em_estro14:*
	em_estro15:*
	em_estro16:*
	em_estro17:*
	em_estro18:*
	em_estro19:*
	em_estro20:*
	gb_est25:*
	gb_est26:*
	gb_est27:*
	gb_est28:*
	gb_est29:*
	gb_est30:*
	gb_est31:*
	gb_est32:*
	gb_est41:*
	gb_est42:*
	gb_est43:*
	gb_est44:*
	gb_est45:*
	gb_est46:*
	gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
C 1	809.2	8.8	912	106	AL582775	AL582775 AL582775
C 2	773.6	8.4	865	106	AL560096	AL560096 AL560096
C 3	688.2	7.5	703	105	AL045547	AL045547 DKFZP434L
C 4	658.2	7.2	898	105	AL521542	AL521542 AL521542
C 5	613.4	6.7	680	165	BE267789	BE267789 601124773
C 6	574.4	6.3	792	106	AL561613	AL561613 601124773
C 7	566.8	6.2	641	122	AW950638	AW950638 EST362708
C 8	560.4	6.1	580	165	BE267618	BE267618 601124837
C 9	532	5.8	574	122	AW968425	AW968425 EST380501
C 10	489.6	5.3	799	150	BF526328	BF526328 602071372
C 11	484	5.3	523	169	BF803247	BF803247 IL5-C1014
C 12	474.4	5.2	876	105	AL521543	AL521543 AL521543
C 13	471.4	5.1	897	106	AL560097	AL560097 AL560097
C 14	466.8	5.1	525	11	AA776738	AA776738 ah49f08.s
C 15	461.8	5.0	525	19	AI333944	AI333944 q27b09.x
C 16	460.8	5.0	533	11	AA743431	AA743431 ny23f01.s
C 17	460.4	5.0	630	165	BE281439	BE281439 601154960
C 18	459	5.0	503	15	AI082242	AI082242 ox79g05.x
C 19	449.4	4.9	497	104	AI953167	AI953167 wp99d02.x
C 20	447.8	4.9	496	116	AW439193	AW439193 xt16c10.x
C 21	445	4.8	490	152	BG328937	BG328937 602428083
C 22	444	4.8	511	10	AA705195	AA705195 zj96e04.s
C 23	440.2	4.8	500	148	BF448417	BF448417 7n8e07.x
C 24	440	4.8	496	145	BF195671	BF195671 7n8e07.x
C 25	440	4.8	507	112	AW150068	AW150068 xg48b02.x
C 26	434.2	4.7	738	175	BG251389	BG251389 602364035
C 27	431	4.7	486	145	BF195544	BF195544 7n89f07.x
C 28	427	4.7	687	152	BG326214	BG326214 602424925
C 29	425.4	4.6	591	112	AW161381	AW161381 aub1a12.y
C 30	422	4.6	467	14	AA973790	AA973790 oa45904.s
C 31	420	4.6	464	22	AI633641	AI633641 th70f01.x
C 32	418.8	4.6	467	191	Z98467	Z98467 HSZ98467 DK
C 33	418.4	4.6	464	18	AI288914	AI288914 q191q06.x
C 34	417.4	4.5	474	17	AI2424561	AI2424561 q134d04.x
C 35	411.4	4.5	856	153	BG422364	BG422364 602446924
C 36	405	4.4	449	18	AI275473	AI275473 q174a11.x
C 37	404	4.4	437	223	AA053443	AA053443 RPC111-51
C 38	402.4	4.4	448	19	AI332627	AI332627 qq28b05.x
C 39	402.4	4.4	460	19	AI371104	AI371104 ta30a02.x
C 40	402	4.4	446	19	AI371103	AI371103 ta30a01.x
C 41	402	4.4	480	9	AA604513	AA604513 no73c10.s
C 42	399.8	4.4	448	10	AA636030	AA636030 nr37h07.s
C 43	394.8	4.3	808	154	BG476481	BG476481 602522015
C 44	394.6	4.3	443	24	AI738671	AI738671 w11g09.x
C 45	392	4.3	437	15	AI056630	AI056630 oz18e01.x

ALIGNMENTS

RESULT 1	AL582775/c	AL582775	912 bp	mRNA	EST	16-FEB-2001
LOCUS	AL582775	LTI_NFL010_BC2	Homo sapiens	cdna clone	CS0DL011Y120	3
DEFINITION	prime, mRNA sequence.					
ACCESSION	AL582775					
VERSION	AL582775.1	G1:12951093				
KEYWORDS	EST.					
SOURCE	human.					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
AUTHORS	Li, W.B., Gruber, C., Jesse, J., and Polayes, D.					
TITLE	Full-length cDNA libraries and normalization					
JOURNAL	Unpublished (2001)					
COMMENT	Contact: Genoscope					

Genoscope - Centre National de Sequencage BP 191 91006 EVRY cedex - France Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.		Location/Qualifiers	
FEATURES		source	
		1..912	
		/organism="Homo sapiens"	
		/db_xref="taxon:9606"	
		/clone="CS0DL011Y120"	
		/clone_lib="LTI_NFL010_BC2"	
		/sex="male"	
		/tissue_type="B cells from Burkitt lymphoma"	
		/notes="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"	
BASE COUNT		179 a 178 c 307 g 214 t 34 others	
ORIGIN			
Query Match		8.8%; Score 809.2; DB 106; Length 912;	
Best Local Similarity		94.5%; Pred. No. 2.3e-64;	
Matches 867; Conservative 29; Mismatches 15; Indels 6; Gaps 6;			
QY	7925	GGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCACTCACGGCACAGTCGCAAGTGCA	7984
DB	912	GGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCACTCACGGCACAGTCGCAAGTGCA	853
QY	7985	GATGAAGAAGCAGAAAGTGCACCCCTAGTACTACACTGTCTTCTTCTCAAGCGCA	8044
DB	852	GATGAAGAAGCAGAAAGTGCACCCCTAGTACTACACTGTCTTCTTCTCAAGCGCA	794
QY	8045	GGCCAAAGGGCTCTGAACTACTTGGGAGCTTTCGGACCGCTTGTGGGGACCCAGGCTCCG-C	8103
DB	793	SCSAAAGSCTCTGAACTACTTGGGAGCTTTCGGACCGCTTGTGGGGACCCAGGCTCCG-C	734
QY	8104	CTTAGTCCCCCACTCTGAGCGCCATGTTCTGCCGCCAGCCCAAGGGGACAGGCTCACC	8163
DB	733	CTTAGTCCCCCACTCTGAGCGCCATGTTCTGCCGCCAGCCCAAGGGGACAGGCTCACC	674
QY	8164	TCATACCAAAACCTTAGTCTCCCGTCCCGAGTACAGTCTGTATCAAAACCCAGATTCT	8223
DB	673	TCATACCAAAACCTTAGTCTCCCGTCCCGAGTACAGTCTGTATCAAAACCCAGATTCT	614
QY	8224	CCAGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTAGTAATATAGGTCTCTTCTCCCAAGAT	8283
DB	613	CCAGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTAGTAATATAGGTCTCTTCTCCCAAGAT	554
QY	8284	CCAGCGGGCCAAATGGAAACCTCAGCTGGGTCTTAATACAGCTCTTTAAAGGCCACG	8343
DB	553	CCAGCGGGCCAAATGGAAACCTCAGCTGGGTCTTAATACAGCTCTTTAAAGGCCACG	494
QY	8344	CCCTAGAAACCCCAAGCTCTCTCCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCA	8403
DB	493	CCCTAGAAACCCCAAGCTCTCTCCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCA	434
QY	8404	GGGCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCCCAAGGAGCAGAGTCC	8463
DB	433	GGGCTCTCCAGCTTGTAGGAGCTGAGGTTTACCCCTTAACCCCAAGGAGCAGAGTCC	375
QY	8464	CACCTCAGCGCCGGGAGCCCTAGGACCACTCAGCCCTAGGAGTATATTTCCGCACTTCA	8523
DB	374	CACCTCAGCGCCGGGAGCCCTAGGACCACTCAGCCCTAGGAGTATATTTCCGCACTTCA	316
QY	8524	GAATTCATCTTTCGGAATCCCAAGCTCCCTGCCGCCCAATTAACCTCACTGCTGCTCCAG	8583
DB	315	GAATTCATCTTTCGGAATCCCAAGCTCCCTGCCGCCCAATTAACCTCACTGCTGCTCCAG	257

```
QY 8584 AATTTGGAATCCTAGTTTCTCTCTCTCGTATCCGAGTCTGGGACACAAACATCCGCC 8643
|||||
Db 256 AATTTGGAATCCTAGTTTCTCTCTCTCGTATCCGAGTCTGGGACAAACATCCMCC 197
|||||
QY 8644 CCACGCTATGAGCATCTGAGCGCCGCCCTCTCTGACGAACTGCGCCGGATCAGA 8703
|||||
Db 196 CCACGCTATGAGCATCTGAGCGCCGCCCTCTCTGACGAACTGCGCCGGATCAGA 137
|||||
QY 8704 GCAGGACCTCCCTCCGACCCCTCTGGAACTCTCCAGAGGTCCAGCCCATCTCGGAGCAT 8763
|||||
Db 136 GCAGGACCTCCCTCCGACCCCTCTGGAACTCTCCAGAGVCCAGCCCATCTCGGAGCAT 77
|||||
QY 8764 CCCGAGGAAATCTCGAGAGGGTTAGGAGTGGGTGACAAAGACCTGTATCTTCTCTGTT 8823
|||||
Db 76 CCCGAGGAAATCTCGAGA-GGGTTAGGAGTGGGTGACAAAGACCTGTATCTTCTCTGTT 18
|||||
QY 8824 TTGTACATAGATTTATT 8840
|||||
Db 17 TAGAACATAGATTTATT 1

RESULT 2
AL560096/C 865 bp mRNA EST 16-FEB-2001
LOCUS AL560096 LTI_FL011_BCl Homo sapiens cDNA clone CS0DG002YG12 3 prime
DEFINITION , mRNA sequence.
ACCESSION AL560096
VERSION AL560096.1 GI:12906228
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 865)
AUTHORS Li W.B., Gruber C., Jessee J. and Polayes D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
Location/Qualifiers
1..865
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DG002YG12"
/clone_lib="LTI_FL011_BCl"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH10B"
/note="vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-stranded cDNA was digested with Not I and cloned
into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
Library was constructed by Life Technologies. Contact :
Feng Liang Life Technologies, a Division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 165 a 183 c 306 g 204 t 7 others
ORIGIN

Query Match 8.4%; Score 773.6; DB 106; Length 865;
Best Local Similarity 97.1%; Pred. No. 3.7e-61;
Matches 845; Conservative 5; Mismatches 14; Indels 6; Gaps 6;

QY 7817 AGTGGCAGCAGGCTCAGGTCCAGCACCCG-CAGCATCACACCGCGGAGGTCCAG 7875
|||||
Db 865 AAGTGCAGCAGGCTCAAGTCCAGCACCCCGCAGCATCACACCGCGGAGGTCCAG 806
|||||
QY 7876 TCCTCACTTCAGAGTGAAGAGTGAAGGATGAAGGAGTCTGTGGCCACCAAGA 7935
|||||
```

```
Db 805 TKCTCACTTTCAGAGTGAGAAAGATGAAGGGATGAAGGAGCTGCT- GTGGCCACCAAGA 747
QY 7936 TCAACTCGAGCGCCCATCAAGCTGCAACTCACGGCACAGTCTGCAAGTGCAGATGCAAGAAGC 7995
|||||
Db 746 TCAACTCGAGCGCCCATCAAGCTGCAACTCACGGCACAGTCTGCAAGTGCAGATGCAAGAAGC 687
|||||
QY 7996 AGAAAGTGTCCACCCCTAGTGTACTACTCTCTCTTCTCAAGCGGCAGCGCAAGGCC 8055
|||||
Db 686 AGAAAGTGTCCACCCCTAGTGTACTACTCTCTCTTCTCAAGCGGCAGCGCAAGGCC 627
|||||
QY 8056 TCTGAACTACTGGGACACTTCGGACCGCTGTGGGACCCAGGCTCGCGCTTCTAGTCCCGCA 8115
|||||
Db 626 TCTGAACTACTGGGACACTTCGGACCGCTGTGGGACCMCCAGTCTCGCGCTTCTAGTCCCGCA 567
|||||
QY 8116 ACTCTGAGCGCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCTCACCTCTACCCAAAGC 8175
|||||
Db 566 ACTCTGAGCGCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCTCACCTCTACCCAAAGC 507
|||||
QY 8176 CTAGTTTCCCGGTCGCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCCAGCTTCAGAAC 8235
|||||
Db 506 CTAGTTTCCCGGTCGCGAGTACAGTCTGTATCAAAACCCACGATTTTCTCCAGCTTCAGAAC 447
|||||
QY 8236 CCAGGCGCTCGCCCCAGTCGTTAGATATAGTCTCTTCTCCAGATCCAGCGGCCA 8295
|||||
Db 446 CCA-GGCTCTGCCCCAGTCGTTAGATATAGTCTCTTCTCCAGATCCAGCGGCCA 388
|||||
QY 8296 ATGGAACCTCACGCTGGGTCTTAATTACAGTCTTTTAAAGGCCACGAGCCCTAGAAACCC 8355
|||||
Db 387 ATGGAACCTCACGCTGGGTCTTAATTACAGTCTTTTAAAGGCCACGAGCCCTAGAAACCC 328
|||||
QY 8356 AAGTCTCTCTCGGAAACCGCTCACCTAGAGCAGACCAACGTTTACTTCAGGGTCTCTCCA 8415
|||||
Db 327 AAGTCTCTCTCGGAAACCGCTCACCTAGAGCAGACCAACGTTTACTTCAGGGTCTCTCCA 268
|||||
QY 8416 GCTTCTAGAGCTGAGGTTTACCCCTTACCCCAAGGGGACAGGTCACACCTCCAGCCC 8475
|||||
Db 267 GCTTCTAGAGCTGAGGTTTACCCCTTACCCCTTACCCCAAGGGGACAGGTCACACCTCCAGCCC 209
|||||
QY 8476 GGGAGCCTTAGGACCACTCAGCCCTCAGGATATATTTCCGACATTCAGAAATTCATATC 8535
|||||
Db 208 -GGGAGCCTTAGGACCACTCAGCCCTCAGGATATATTTCCGACATTCAGAAATTCATATC 150
|||||
QY 8536 TTGCGAATCCAAAGTCCCTGCCCCCAATAAATTCAGTCTCTGTTCCAGAAATTTGGAATC 8595
|||||
Db 149 TTGCGAATCCAAAGTCCCTGCCCCCAATAAATTCAGTCTCTGTTCCAGAAATTTGGAATC 91
|||||
QY 8596 CTAGTTTCTCTCTCTCTCTATCCGAGTCTGGGACACAAACTCCGCCCGGAGCTTATCA 8655
|||||
Db 90 CTAGTTTCTCTCTCTCTCTATCCGAGTCTGGGACACAAACTCCGCCCGGAGCTTATCA 31
|||||
QY 8656 GCATCCTGAGCCCGCCCTCTTCTGACGA 8685
|||||
Db 30 GCATCCTGAGCCCGCCCGCCCTCTACGA 1
|||||

RESULT 3
AL045547 703 bp mRNA EST 29-FEB-2000
LOCUS AL045547
DEFINITION DKEZp434L125_r1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045547
VERSION AL045547.1 GI:5433678
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 703)
AUTHORS Duesterhoeft A., Lauber J., Mewes H.W., Gassenhuber J. and Wiemann S.
TITLE EST (Duesterhoeft, et al.)
JOURNAL Unpublished (1999)
COMMENT Contact: Duesterhoeft A
```

MIPS

Am Klopferspitz 18a D-82152 Martinsried, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKF2); Email s.wiemann@dkfz-heidelberg.de;
consortium of Otagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
NO 5' sequence available.

This clone (DKF2p34L125) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES

Location/Qualifiers
1..703
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="DKF2p34L125"
/clone_lib="434 (synonym: htes3)"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/note="Vector: pSport1; Site_1: NotI; Site_2: SalI"
136 a 227 c 185 g 151 t 4 others

BASE COUNT 136 a 227 c 185 g 151 t 4 others
ORIGIN
Query Match 7.5%; Score 688; DB 105; Length 703;
Best Local Similarity 99.4%; Pred. No. 1.9e-53;
Matches 699; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

Qy 5720 GCCATTGGCTGGCTGGAAGTCTTTGGTGTAGTGTAGAGAGTGTCTGAGAAGAGAA 5779
Db 1 GCCATTGGCTGGCTGGAAGTCTTTGGTGTAGTGTAGAGAGTGTCTGAGAAGAGAA 60
Qy 5780 GGGCCCTGAGCTCGAGGGCAGGCCACCCCTCGAGTCTGCCCCAGGCTCAGCCAGCA 5839
Db 61 GGGCCCTGAGCTCGAGGGCAGGCCACCCCTCGAGTCTGCCCCAGGCTCAGCCAGCA 120
Qy 5840 GTCTGTAGACCCAGGAGGAGACAGGTAGAAGGGCTGGCAGCAGTGGAGTGGGAGT 5899
Db 121 GTCTGTAGACCCAGGAGGAGACAGGTAGAAGGGCTGGCAGCAGTGGAGTGGGAGT 180
Qy 5900 GGAGATGGAGAGACTCCCTGGGATCTTCCTGTGGCCCTTCTGGGTGGCCCTGGTGGG 5959
Db 181 GGAGATGGAGAGACTCCCTGGGATCTTCCTGTGGCCCTTCTGGGTGGCCCTGGTGGG 240
Qy 5960 GCATTGTGCGCAGGAGGAGCTGGGCTGCTCCCTGAGGATCCTGCTCCTCACTCCA 6019
Db 241 GCATTGTGCGCAGGAGGAGCTGGGCTGCTCCCTGAGGATCCTGCTCCTCACTCCA 300
Qy 6020 TCCAGGCAATTGCTCAGCCAGACCTACTATCGGGATGAACACATCTACCCCTACATGT 6079
Db 301 TCCAGGCAATTGCTCAGCCAGACCTACTATCGGGATGAACACATCTACCCCTACATGT 360
Qy 6080 ACCTGGCTGGCTACACTGTGCAACCGCAATGTGGGGAAGCCCTCGAGGCCCTGGGGG 6139
Db 361 ACCTGGCTGGCTACACTGTGCAACCGCAATGTGGGGAAGCCCTCGAGGCCCTGGGGG 420
Qy 6140 ACAGGGCACTGTCTATCCAGGAGTGGATCCGCCCTACTAGGGCTCGCAGCCTGTCTTT 6199
Db 421 ACAGGGCACTGTCTATCCAGGAGTGGATCCGCCCTACTAGGGCTCGCAGCCTGTCTTT 480
Qy 6200 CTTCCCTTCCATCAGTTTCCAAACCACTCGTCCAGGACTGAGGCCCTGGCTCCACGCC 6259
Db 481 CTTCCCTTCCATCAGTTTCCAAACCACTCGTCCAGGACTGAGGCCCTGGCTCCACGCC 540
Qy 6260 CATCCCTTTCCATCCAGTCCCTAGGAGCAAGGCCACCATTTACCA-GGAGTAGGGAC 6318
Db 541 CATCCCTTTCCATCCAGTCCCTAGGAGCAAGGCCACCATTTACCAAGGAGTAGGGAC 600
Qy 6319 CCTGATTAAGGTGTACATCTTCCCTCCCTCCCTCCCTCCCTAAATTTTTTTTCTCA 6378
Db 601 CCTGATTAAGGTGTACATCTTCCCTCCCTCCCTCCCTCCCTAAATTTTTTTTCTCA 660

Qy 6379

GAACAGTCTCAATCTCAATGTTTAAACCACCATCATCCAGCA 6421
|||||
Db 661 GAACAGTCTCAATCTCAATGTTTAAACCACCATCATCCANCA 703

RESULT 4

AL521542/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

CONTACT

Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1..898

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CS0DB003YA04"

/clone_lib="LTI_NFL004_NBC2"

/sex="male"

/tissue_type="neuroblastoma cells"

/lab_host="DH10B"

/notes="Organ: Brain; Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com"

BASE COUNT 190 a 188 c 297 g 206 t 17 others

ORIGIN

Query Match 7.2%; Score 658.2; DB 105; Length 898;

Best Local Similarity 89.0%; Pred. No. 7.6e-51;

Matches 803; Conservative 8; Mismatches 68; Indels 23; Gaps 9;

Qy 7978 AAGTGCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTACTACTACTCTGTCTTCTCA 8037

Db 898 AAGTGSAGATGAAGAGCAGAAAGTGTCCACCCCTAGTACTACTACTCTGTCTTCTCA 839

Qy 8038 AGCGGCGAGCAAGGCTCTGAACACTCTGGGAGCTTCGGACCGCTTGTGGGACCCAG 8097

Db 838 AGCGGCGAGCAAGGCTCTGAACACTCTGGGAGCTTCGGACCGCTTGTGGGACCCAG 779

Qy 8098 CTCGCCCTTAGTCCGCCCAACTCTGAGCCCATGTCTTGTCCGCCCAAGGAGGAGG 8157

Db 778 CTCGCCCTTAGTCCGCCCAACTCTGAGCCCATGTCTTGTCCGCCCAAGGAGGAGG 719

Qy 8158 CTCACCTTACCCAAACCTAGGTTCGGGTCCCGAGTACAGTCTGTATCAACCCAGCA 8217

Db 718 CTCACCTTACCCAAACCTAGGTTCGGGTCCCGAGTACAGTCTGTATCAACCCAGCA 659

Qy 8218 TTTTCTCCAGTCTAGACCCAGGCTCTGCCCCAGTCCGTTAGATATAGTCTCTTCTCC 8277

Db 658 TTTTCTCCAGTCTAGACCCAGGCTCTGCCCCAGTCCGTTAGATATAGTCTCTTCTCC 600

Qy 8278 CAGATCCAGCCGCGCAATGGAAACCTCAGGCTGGGTCTTAATTTACCAGTCTTTAAAGG 8337

DEFINITION AL561613 LTI_NFL010_BC2 Homo sapiens cDNA clone CS0DL011Y120 5
prime, mRNA sequence.
ACCESSION AL561613
VERSION AL561613.1 GI:12909216
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 792)
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
source
1..792
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DL011Y120"
/clone_lib="LTI_NFL010_BC2"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/note="Vector: pCMVSPORT 6; Site 1: NotI; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-stranded cDNA was digested with NotI and cloned into the NotI and EcoRV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371. Email : fliang@lifetech.com URL : http://fulllength.invitrogen.com."

BASE COUNT 119 a 259 c 252 g 157 t 5 others
ORIGIN

Query Match 6.3%; Score 574.4; DB 106; Length 792;
Best Local Similarity 84.0%; Pred. No. 2.6e-43;
Matches 739; Conservative 4; Mismatches 0; Indels 137; Gaps 3;

QY 1854 GGTCCTAGTGTGGATGTAAGCGCGAGGTGGCGAGGGGACCGAGGCGAGGACTCTC 1913
Db 1 GGTCCCTAGTGTGGATGTAAGCGCGAGGTGGCGAGGGGACCGAGGCGAGGACTCTC 60
QY 1914 CTTGGGTTTGGGGCTTGACCTGGTGGCGCTTTCTGGACAGACTTTACAGCCCCGGGG 1973
Db 61 CTTGGGTTTGGGGCTTGACCTGGTGGCGCTTTCTGGACAGACTTTACAG-CCCCGGGG 119
QY 1974 GCACAGTCGTAGAGAGGGGGGGCGCCATTGGGGCTCTCATTTGGGTGCTTGGGGC 2033
Db 120 GCACAGTCGTAGAGAGGGGGGGGGCGCCATTGGGGCTCTCATTTGGGTGCTTGGGGC 179
QY 2034 GCACCCCATCGGTACCGGGCGTCCCGGAATTGTGGGGGACAAAAGGCTCTGCAGTCTC 2093
Db 180 GCACCCCATCGGTACCGGGCGTCCCGGAATTGTGGGGGACAAAAGGCTCTGCAGTCTC 239
QY 2094 GGCTGAGGGTCTCACCACACAAAGAGGGGAGCCGGTGTAGCAGAGGCTGAAGAGGGTGG 2153
Db 240 GGCTGAGGGTCTCACCACACAAAGAGGGGAGCC----- 274
QY 2154 GGAAGCAGGGGAGGCTGTGGCTGTGTCGGGGGGGTGGACCTTACCGACCTCGGAGGA 2213
Db 275 ----- 274
QY 2214 GGCTCCCGCGCGGAACCTGCCCGACCTCCCTCCCGGCTTGCTTGCAGGGCGCGGCC 2273
Db 275 -----GGCCCGCGGCC 284
QY 2274 CACCGCCCGCGGCATGGGGCTGAAGCCCGCCAGAGAGCGCTGTCCCGCTGCCTCCA 2333
|||||

Db 285 CACGCCCGCGCCATGGGCTGAAGCGCGCCAGAAAGACGCTGTCCCGCTGGCTCCA 344
QY 2334 TCAGACAGCTGGTGGCGCTGTTCCTCCGAGCTGGCGGAGAGAGAGCGGACCTGGTC 2393
Db 345 TCAGACAGCTGGTGGCGCTKTTTCTCCGAGCTGGCGGAGAGAGCGGACCTGGTC 404
QY 2394 TCCTTTCTTGGTGGCTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTA 2453
Db 405 TCCTTTCTTGGTGGCTGGGCTTCGTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTA 464
QY 2454 CCACAGTTCGGAGCTCACCTTCCAGCCGAGCCCGCCGAGCCGCTGGCGGCTCA 2513
Db 465 CCAAGTTCGGAGCTCACCTTCCAGCCGAGCCCGCCGAGCCGCTGGC-GCCTCA 523
QY 2514 CCTACTTTCCGTCGGCGAGCTGTCTATCATCGCGCGCTCTATGCGCGCTTCAACCGCC 2573
Db 524 CCTACTTTCCGTCGGCGAGCTGTCTATCATCGCGCGCTCTATGCGCGCTTCAACCGCC 583
QY 2574 AGATCCGAGCGCGCTGCAGCTGTCCCTCTATCTCGAGAGAGGGGTGTCTCCAGCGCTG 2633
Db 584 AGATCCGAGCGCGCTGCAGCTGTCCCTCTATCTCGAGAGAGGGGTGTCTCCAGCGCTG 643
QY 2634 AGCTGGTGAAGAGGTCTCGATGTCATATGGAACAGCTCAGCCGCTCTCTACTTCAAGG 2693
Db 644 AGCTGGTGAAGAGGTCTCGATGTCATATGGAACAGCTCAGCCGCTCTCTACTTCAAGG 703
QY 2694 ATCGGGCCCATCATCCAGTCCCTCTTCACTTCACTACACAGG 2733
Db 704 ATCGGGCCCATCATCCAGTCCCTCTTCACTTCACTACACAGG 743

RESULT 7
AW950638 641 bp mRNA EST 01-JUN-2000
LOCUS EST382708 MAGE resequences, MAGA Homo sapiens cDNA, mRNA sequence.
ACCESSION AW950638
VERSION AW950638.1 GI:8140296
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 641)
AUTHORS Hegde, P., Qi, R., Abernathy, K., Dharap, S., Gaspar, R., Gay, C., Holt
I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and
Quackenbush, J.
TITLE Assessment of gene expression patterns in a model of colon tumor
metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)
COMMENT Contact: John Quackenbush
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: johnd@tigr.org
Plate: 17
Seq primer: Reverse.
Location/Qualifiers
1..641
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="MAGE resequences, MAGA"
/note="Vector: pBluescriptSKm"

BASE COUNT 152 a 214 c 156 g 119 t
ORIGIN

Query Match 6.2%; Score 566.8; DB 122; Length 641;
Best Local Similarity 96.0%; Pred. No. 1.4e-42;
Matches 603; Conservative 0; Mismatches 22; Indels 3; Gaps 2;

QY 7765 GACCCCGCGGAGGCTCTCTGGGACTGTGCTGGCAGACCGCGGCTTCAAGTGSCA 7824
|||||

Authors: Heyde, P., Qi, R., Abernathy, K., Dharap, S., Gaspard, R., Gay, C., Holt,
 I. E., Saeed, A. I., Sharov, V., Lee, N. H., Yeatman, T. J. and
 Quackenbush, J.
 Title: Assessment of gene expression patterns in a model of colon tumor
 metastasis using a 19,200 element cDNA microarray
 Journal: Unpublished (2000)
 Comment: Contact: John Quackenbush
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 3528
 Fax: 301 838 0208
 Email: johnq@tigr.org
 Plate: 253
 Seq primer: Forward.
 Location/Qualifiers
 1. 574
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resources, MAGJ"
 /note="vector: pbluescriptSKm"
 133 a 205 c 117 g 119 t
 BASE COUNT
 ORIGIN

	Query Match	5.8%; Score 532; DB 122; Length 574;
	Best Local Similarity	98.6%; Pred. No. 1.9e-39;
	Matches 568; Conservative	0; Mismatches 5; Indels 3; Gaps 3;
Qy	8217	ATTTTCTCCAGCTCAGAACCCAGGCGCTCTGCCCAAGTCGTTAGAATAATAGTGCTCTTCTC 8276
Dy		
Dy	1	ATTTTCTCCAGCTCAGAACCCAGGCGCTCTGCCCAAGTCGTTAGAATAATAGTGCTCTTCTC 60
Qy	8277	CCAGAATCCCAGCGGGCCAATGGAAACCTCACGTGGGTCTTAATTACCAGCTTTTAAAG 8336
Dy		
Dy	61	CCAGAATCCCAGCGGGCCAATGGAAACCTCACGTGGGTCTTAATTACCAGCTTTTAAAG 120
Qy	8337	GCCAGCGCCCTAGAAACCCAGACTCTCTCGGAACCCGCTCACCTAGAGCCAGACCAAG 8396
Dy		
Dy	121	GCCAGCGCCCTAGAAACCCAGACTCTCTCGGGGAACCGCTCACCTAGAGCCAGACCAAG 180
Qy	8397	TTACTCAGGGGCTCTCCAGAGCTGTAGGAGCTGAGGTTTTACCCCTTAAACCAAAGGAGCA 8456
Dy		
Dy	181	TTACTCAGGGGCTCTCCAGAGCTGTAGGAGCTGAGGTTTTACCCCTTAAACCAA - GGAGCA 239
Qy	8457	CAGGTCCCACTCCAGCCCCGGGAGCCTTAGGACCACCTCAGCCCCCTAGGAGTATATTTCCG 8516
Dy		
Dy	240	CAGGTCCCACTCCAGCCC - GGGAGCCTTAGGACCACCTCAGCCCCCTAGGAGTATATTTCCG 298
Qy	8517	CACTTTCAGAAATTCATATCTTTCGGAATCCAGCTCCCTGCCTCCCAATAACTTCAGTCTCTG 8576
Dy		
Dy	299	CACTTTCAGAAATTCATATCTTTCGGAATCCAGCTCCCTGCCTCCCAATAACTTCAGTCTCTG 358
Qy	8577	CTTCAGAAATTTGAAAATCCCTAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8636
Dy		
Dy	359	C - TCCAGAAATTTGAAAATCCCTAGTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 417
Qy	8637	CTCGCCCCCAGCCTATGAGCATCTGAGCCCCGCCCTCTTCTCTGACAAACTGCCCCCG 8696
Dy		
Dy	418	CTCGCCCCCAGCCTATGAGCATCTGAGCCCCGCCCTCTTCTCTGACGAACTTGCCCCCG 477
Qy	8697	GATCAGACGAGGACCTCCCTTTCCGACCCTCTGGGAACCTCCAGAGGTCCAGCCCCATCTC 8756
Dy		
Dy	478	GATCAGACGAGGACCTCCCTTTCCGACCCTCTGGGAACCTCCAGAGGTCCAGCCCCATCTC 537
Qy	8757	GGAGCATCCGGAGGAATCTGCAGAGGGGTTAGGA 8792
Dy		
Dy	538	GGAGCATCCGGAGGAATCTGCCAAGGGTTAGAA 573

RESULT 10					
BF526328					
LOCUS	BF526328	799 bp	mrna	EST	11-DEC-2000
DEFINITION	60207137F1	NCI CGAP Brn64	Homo sapiens	cDNA	cloneIMAGE:4214272

```

5', mRNA sequence.
BF526328
VERSION BF526328.1 GI:11613679
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 799)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: ccrabps-remail.nih.gov
Tissue Procurement: David N. Louis, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM9787 row: p column: 17
High quality sequence start: 3
High quality sequence stop: 644.
Location/Qualifiers
1..799
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4214272"
/clone_lib="NCI_CGAP_Brn64"
/tissue_type="glioblastoma with EGFR amplification"
/lab_host="DH10B (T1 phage-resistant)"
/note="Organ: brain; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.57 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
185 a 243 C 266 g 104 t 1 others
BASE COUNT
ORIGIN

```

Query Match	5.3%	Score 489.6	DB 150	Length 799	
Best Local Similarity	87.3%	Pred. No. 9.9e-36			
Matches	559	Conservative 0	Mismatches 79	Indels 2	Gaps 2
Qy 7499	GGATGGCCAGAGCAGGCTCTGGAGTTCCACCCACTGGCCGGCAACCTTGCTCTCACCTT	7558			
Db 83	GGAGGCACTCCACGCTGTGCTGCAGTGGCTGGCCACCTTCTTGTGCAGTCCT	142			
Qy 7559	GCTCTCCCCACTGGCCCCAGGTGCGGCAGAGAGTGCGCATATGTAGCCGAGAGGCCGAGGC	7618			
Db 143	AGGCGCTTTTGGAGGACAGGTGCGGCAGAAAGTGCATAGTAGCCGAGAGGCGGAGGC	202			
Qy 7619	GGCCAGGCCAGGAGCGCTGGGCGAGGAGCCGGGAAGCGCGGCGGGGGGCCACG	7678			
Db 203	GGCCAGGCCAGAGAGCCCTGGGCGAGGAAGCCGGGAAGCCGGCGGGGGCCACG	262			
Qy 7679	CGGGAGTTCGAAGCCAGAGAGCCCGCCGCCAAGAGCAGCACTGGACAAGGGCCT	7738			
Db 263	CGGGAGTTCGAAGCCAGAGAGCCCGCCGCCAAGAGCAGCACTGGACAAGGGCCT	322			
Qy 7739	GGGCACCGCCAGAGGTGCAGTGTTCAGGACCCCGCGAAGCTCTCTGGCACTGTCGTGG	7798			
Db 323	GGGCACCGCCAGAGGTGCAGTGTTCAGGACCCCGCGAAGCTCTCTGGCACTGTCGTGG	382			
Qy 7799	CACAGCCCGAGGCCCTGAAGTGGCAGCACGGCTCAGTGTCCAGCACCCCGCAGCATACC	7858			
Db 383	CACAGCCCGAGGCCCTGAAGTGGCAGCACGGCTCAGTGTCCAGCACCCCGCAGCATACC	442			
Qy 7859	ACCGCGGAGGTCCAGTGCTCACTTTCCAGAGTGAGNAGATGAAGGCATGAAGAGCT	7918			
Db 443	ACCGCGGAGGTCCAGTGCTCACTTTCCAGAGTGAGNAGATGAAGGCATGAAGAGCT	502			
Qy 7919	GCTGTTGCCCAACGAATCAACTCGAGGCGCATCAAGCTGCAACTCAGCGGCACAGTCGCA	7978			

```

|||||
Db 503 GCTGTGGCCACCAAGATCAAGTCGAGGCCATCAAGTCGAACTCAGCGGACAGTCGCA 562
|||||
Qy 7979 AGTCAGATGAAGAG-CAGAAAGTCCACCCCTAGTACTACACTCTGTCTTCTCTCA 8037
|||||
Db 563 AGTCAGATGACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 622
|||||
Qy 8038 AGCGGACGCAAGGCGCTCTGAACCTACTGGGAGCTTCGGACCGCTTGTGGGACCCAGG 8097
|||||
Db 623 AAGCGGAGCAAGGCGCTCTGAACCTACTGGGAGCTTCGGACCGCTTGTGGGAG-CCAGG 681
|||||
Qy 8098 CTCGCCCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCC 8137
|||||
Db 682 ATCCGCCTAAGTCCCGCAACTCTGAGCCCATGTCGAGGCC 721
|||||

RESULT 11
LOCUS BF803247/c 523 bp mRNA EST 12-JAN-2001
DEFINITION IL5-C10149-011100-224-g09 C10149 Homo sapiens cDNA, mRNA sequence.
ACCESSION BF803247
VERSION BF803247.1 GI:12132236
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 523)
Dias Neto, E., Garcia Correa, R., Verjovski-Almeida, S., Briones, M.R.,
Negai, M.A., da Silva, W. Jr., Zago, M.A., Bordin, S., Costa, F.F.,
Goldman, G.H., Carvalho, A.F., Matsukuma, A., Baia, G.S., Simpson, D.H.,
Brunstein, A., de Oliveira, P.S., Bucher, P., Jongeneel, C.V., O'Hare,
M.J., Soares, F., Brentani, R.R., Reis, L.F., de Souza, S.J. and
Simpson, A.J.
Shotgun sequencing of the human transcriptome with ORF expressed
sequence tags
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
MEDLINE 20202663
COMMENT Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research
Rua Prof. Antonio Prudente 109, 4 andar, 01509-010, Sao Paulo-SP,
Brazil
Tel: +55-11-2704922
Fax: +55-11-2707001
Email: asimpson@ludwig.org.br
This sequence was derived from the FAPESP/LICR Human Cancer Genome
Project. This entry can be seen in the following URL
(http://www.ludwig.org.br/scripts/gethtml2.pl?tl=IL5&t2=IL5-C10149-
011100-224-g09&t3=2000-11-01&t4=1)
Seq primer: puc 18 forward
High quality sequence stop: 398.
FEATURES
Location/Qualifiers
source
1..523
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="C10149"
/dev_stage="Adult"
/note="Organ: colon_ins; Vector: puc18; Site_1: SmaI;
Site_2: SmaI; A mini-library was made by cloning products
derived from ORESTES PCR (U.S. Letters Patent application
No. 196,716 - Ludwig Institute for Cancer Research)
profiles into the pUC 18 vector. Reverse transcription of
tissue mRNA and cDNA amplification were performed under
low stringency conditions."
BASE COUNT 107 a 107 c 189 g 118 t 2 others
ORIGIN
Query Match 5.3%; Score 484; DB 169; Length 523;
Best Local Similarity 99.0%; Pred. No. 4e-35; Indels 3; Gaps 3;
Matches 517; Conservative 0; Mismatches 2;

```

```

Qy 8054 CCTCTGAACCTACTGGGACTTTCGGACCGCTTGTGGGACCCAGGCTCG-CCTTAGTCC 8112
|||||
Db 523 CCTCTGAACCTACTGGGACTTTCGGACCGCTTGTGGGACCCAGGCTCGCGCCCTTAGTCC 464
|||||
Qy 8113 CCAACTCTGAGCCCATGTTCTGCCCCCAGCCCAAGGGAGGAGGCGCTCACCCTCTACCCAA 8172
|||||
Db 463 CCAACTCTGAGCCCATGTTCTGCCNCCAGCCCAAGGGAGGAGGCGCTCACCCTCTACCCAA 404
|||||
Qy 8173 ACCCTAGTCTCCCGGTCGGAGTACAGTCTGTATCAAAACCACCATTTTCTCCAGCTCAG 8232
|||||
Db 403 ACCCTAGTCTCCCGGTCGGAGTACAGTCTGTATCAAAACCACCATTTTCTCCAGCTCAG 344
|||||
Qy 8233 AACCCAGGCGTCTGCCCGCAGTCTGTAGATAATAGTCTCTTCTCCAGAAATCCAGCGG 8292
|||||
Db 343 AACCCAGGCGTCTGCCCGCAGTCTGTAGATAATAGTCTCTTCTCCAGAAATCCAGCGG 284
|||||
Qy 8293 CCAATGGAACCTCAGCGTGGTCTCTTAATACCACTCTTTTAAAGGCCAGGCCCTTAGAAA 8352
|||||
Db 283 CCAATGGAACCTCAGCGTGGTCTCTTAATACCACTCTTTTAAAGGCCAGGCCCTTAGAAA 224
|||||
Qy 8353 CCCAAGCTCCTCTCGGAACCGCTCACCTAGACGACAGCAACGTTACTCAGGGCTCCCTC 8412
|||||
Db 223 CCCAAGCTCCTCTCGGAACCGCTCACCTAGACGACAGCAACGTTACTCAGGGCTCCCTC 164
|||||
Qy 8413 CCAGCTTGTAGAGCTGAGGTTTCCACCTTAAACCAAGGGAGACAGGTCCACCTCCAG 8472
|||||
Db 163 CCAGCTTGTAGAGCTGAGGTTTCCACCTTAAACCAAGGGAGACAGGTCCACCTCCAG 105
|||||
Qy 8473 CCCGGGAGCTAGGACCACTCAGCCCTTAGAGTATATTTCCGCACTTCAGAAATTCAT 8532
|||||
Db 104 CCC-GGGAGCGCTAGGACCACTCAGCCCTTAGAGTATATTTCCGCACTTCAGAAATTCAT 46
|||||
Qy 8533 ATCTTGCAATCCAAGCTCCCTCGCCCAATAACTTTCAGTCC 8574
|||||
Db 45 ATCTTGCAATCCAAGCTCCCTCGCCCAATAACTTTCAGTCC 4
|||||

RESULT 12
LOCUS AL521543 876 bp mRNA EST 13-FEB-2001
DEFINITION AL521543 LTI_NFL004_NBC2 Homo sapiens cDNA clone CS0DB003YA04 5
prime, mRNA sequence.
ACCESSION AL521543
VERSION AL521543.1 GI:12785036
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 876)
Li, W.B., Gruber, C., Jesse, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
Location/Qualifiers
source
1..876
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="CS0DB003YA04"
/clone_lib="LTI_NFL004_NBC2"
/sex="male"
/tissue_type="neuroblastoma cells"
/note="Organ: brain; Vector: pCMVSPORT 6; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed
by Life Technologies. Contact : Feng Liang Life
Technologies, a division of Invitrogen 9800 Medical Center

```

Drive Rockville, Maryland 20850, USA Fax : (1) 301 610
8371 Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com
BASE COUNT 153 a 281 c 256 g 180 t 6 others
ORIGIN

Query Match 5.2%; Score 474.4; DB 105; Length 876;
Best Local Similarity 98.5%; Pred. No. 2.1e-34;
Matches 475; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

QY 2267 CGCGCCACCCCGCCGATGGGCTGAAGCGGCCGAGACGCTGTTCCGGCTG 2326
Db 1 CGCGCCACCCCGCCGATGGGCTGAAGCGGCCGAGACGCTGTTCCGGCTG 60
QY 2327 CGCTCATCGACGCTGTGGCTGTTTCTGCGGAGCTGGCGGAGAGCGCGAC 2386
Db 61 CGCTCATCGACGCTGTGGCGCTGTTTCTGCGGAGCTGGCGGAGAGCGCGAC 120
QY 2387 CTGGTGCTCTTCTTCTGCTGGGCTTCGTGGAGCATTTCTGGCTGTCAACCGCGTC 2446
Db 121 CTGGTGCTCTTCTTCTGCTGGGCTTCGTGGAGCATTTCTGGCTGTCAACCGCGTC 180
QY 2447 ATCCCTACCAAGTTCCTCCGAGCTCACTTCCAGCCGAGCGCCGCGCGCTGCG 2506
Db 181 ATCCCTACCAAGTTCCTCCGAGCTCACTTCCAGCCGAGCGCGCGCGCTGCG 240
QY 2507 GGCCTACCTACTTCCCGTGGCGGACCTGTCTATCATCGCGCCCTCTATGCCGCTTC 2566
Db 241 GGCCTACCTACTTCCCGTGGCGGACCTGTCTATMATCGCGCCCTCTATGCCGCTTC 300
QY 2567 ACCGCCAGATCCGAGCGCCGCTCGACCTGTCCCTCTATCTCTCGAAGGGGGTGTCTCC 2626
Db 301 ACCGCCAGATCCGAGCGCGCTCGACCTGTCCCTCTATCTCTCGAAGGGGGTGTCTCC 360
QY 2627 ACCGTGAGCTGGTGAAGAGTCTCCGATGTCATATGAAGAGCTCAGCGCTCTCTAC 2686
Db 361 AGCCGTGAGCTGGTGAAGAGTCTCCGATGTCATATGAAGAGCTCAGCGCTCTCTAC 420
QY 2687 TTCAAGGATCGGGCCACATCCAGTCCCTCTTCAGCTTCATCAGAGTGGAGCCAGTA 2746
Db 421 TTCAAGGATCGGGCCAAATCCAGTCCCTCTTCAGCTTCATCAGAGTGGAGCCAGTA 480
QY 2747 GG 2748
Db 481 GG 482

RESULT 13
AL560097 897 bp mRNA EST 16-FEB-2001
LOCUS AL560097 LTI_FL011_BCl Homo sapiens cDNA clone CS0DG002YG12 5 prime
DEFINITION mRNA sequence.

ACCESSION AL560097
VERSION AL560097.1 GI:12906230
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 897)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

FEATURES
1..897
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CS0DG002YG12"

/clone_lib="LTI_FL011_BCl"
/sex="male"
/tissue_type="B cells from Burkitt lymphoma"
/lab_host="DH108"
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-stranded cDNA was digested with Not I and cloned
into the Not I and Eco RV sites of the pCMVSPORT 6 vector.
Library was constructed by Life Technologies. Contact :
Peng Liang Life Technologies, a division of Invitrogen
9800 Medical Center Drive Rockville, Maryland 20850, USA
Fax : (1) 301 610 8371 Email : filiang@lifetech.com URL :
http://fulllength.invitrogen.com"
BASE COUNT 152 a 291 c 268 g 182 t 4 others
ORIGIN

Query Match 5.1%; Score 471.4; DB 106; Length 897;
Best Local Similarity 99.8%; Pred. No. 3.9e-34;
Matches 472; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2261 GCAGGCGCGCGCCGCGCCATGGGGCTGAAGCGGCCGAGACGCTGTTC 2320
Db 13 GCAGGCGCGCGCCGCGCCATGGGGCTGAAGCGGCCGAGACGCTGTTC 72
QY 2321 CGCTGCGCTCCATCGACGAGCTGGTGGCGCTGTTTGTCTGCCGAGCTGGCGCGAGGAG 2380
Db 73 CGCTGCGCTCCATCGACGAGCTGGTGGCGCTGTTTGTCTGCCGAGGAG 132
QY 2381 CGGAGCTGGTGTCTCTTCTTGGTGGGCTTCGTGGAGCATTTCTGGCTGTCAAC 2440
Db 133 CGGAGCTGGTGTCTCTTCTTGGTGGGCTTCGTGGAGCATTTCTGGCTGTCAAC 192
QY 2441 CGGCTATCCCTACCAACGTTCCGAGCTCACCTTCCAGCCGAGCGCCGCGCCGCGCG 2500
Db 193 CGGCTATCCCTACCAACGTTCCGAGCTCACCTTCCAGCCGAGCGCCGCGCCGCGCG 252
QY 2501 CTGCGCGGCTCACCTACTTTCCTGGCGGCTGTCTATCATCGCGCCCTCTATGCC 2560
Db 253 CTGCGCGGCTCACCTACTTTCCTGGCGGCTGTCTATCATCGCGCCCTCTATGCC 312
QY 2561 CGCTTCCAGCGCCAGATCCGAGCGCGCTGCGACCTGTCCCTCTATCTTCGAGAGGGGT 2620
Db 313 CGCTTCCAGCGCCAGATCCGAGCGCGCTGCGACCTGTCCCTCTATCTTCGAGAGGGGT 372
QY 2621 GTCTCCAGCGCTCAGCTGGTGAAGAGGCTCCGATGTCATATGAAGACGCTCAGCGCG 2680
Db 373 GTCTCCAGCGCTCAGCTGGTGAAGAGGCTCCGATGTCATATGAAGACGCTCAGCGCG 432
QY 2681 TCCTACTTCAAGGATCGGGCCACATCCAGTCCCTCTTTCAGCTTCATCAGG 2733
Db 433 TCCTACTTCAAGGATCGGGCCACATCCAGTCCCTCTTTCAGCTTCATCAGG 485

RESULT 14
AA776738/c
LOCUS AA776738
DEFINITION mRNA sequence.
ACCESSION AA776738
VERSION AA776738.1 GI:2836072
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 525)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov

cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima Bonaide
Ph.D.

cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1019 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 486.

FEATURES

source
Location/Qualifiers
1. 525
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="1292871"
/clone_lib="Soares_testis_NHT"
/sex="male"
/lab_host="DH10B"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
was prepared from mRNA obtained from Clontech Laboratories
, Inc., and primed with a Not I - oligo(dT) primer [5',
TGTTACCACTCAAGTGGAGCGCGGCCCAATTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaide."

BASE COUNT
ORIGIN

Query Match 5.1%; Score 466.8; DB 11; Length 525;
Best Local Similarity 98.7%; Pred. No. 1.4e-33;
Matches 523; Conservative 0; Mismatches 2; Indels 5; Gaps 5;
Qy 8347 TAGAACCAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCAGG 8406
Db 525 TAGAACCAGCTCTCTCGGAACGCTCACCTAGAGCCAGACCAACGTTACTCAGG 466
Qy 8407 CTCCTCCAGCTTCTAGAGCTGAGGTTTACCCCTTACCCAAAGGACAGCTCCAC 8466
Db 465 CTCCTCCAGCTTCTAGAGCTGA-GTTTACCCCTTACCCAA-GGAGCAGCTCCAC 408
Qy 8467 CTCAGCCCGGGAGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACTTCAGAA 8526
Db 407 CTCAGCCCGGGAGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACTTCAGAA 349
Qy 8527 TTCATATCTTTGCGAATCCAGCTCCCTGCCCCAAATAAATTCAGTCTGCTTCAGAA 8586
Db 348 TTCATATCTTTGCGAATCCAGCTCCCTGCCCCAAATAAATTCAGTCTGCTTCAGAA 290
Qy 8587 TTGGAATCTAGTTCTCTCTCTGTTATCCGAGTCTGGGACACAAACTCCGCCCC 8646
Db 289 TTGGAATCTAGTTCTCTCTCTGTTATCCGAGTCTGGGACACAAACTCCGCCCC 230
Qy 8647 AGCCTATGAGCATCTGAGCCCGCCCTCTTCTTCCGAGAACTGGCCCGGATCAGAC 8706
Db 229 AGCCTATGAGCATCTGAGCCCGCCCTCTTCTTCCGAGAACTGGCCCGGATCAGAC 170
Qy 8707 GGACTCTCCCTTCCGAGCTCTGGAACCTCCAGAGTCCAGCCATCTCGGAGATCCC 8766
Db 169 GGACTCTCCCTTCCGAGCTCTGGAACCTCCAGAGTCCAGCCATCTCGGAGATCCC 110
Qy 8767 GGAGAAATCTGCAGAGGGTTAGAGTGGGTGACAGAGCTGATCTCTTCTGTTTG 8826
Db 109 GGAGAAATCTGCAGA-GGGTTAGAGTGGGTGACAGAGCTGATCTCTTCTGTTTG 51
Qy 8827 TACATAGATTTATTTTTCAGTTTCAAGAAGATGAATACATTTTGTATAA 8876
Db 50 TACATAGATTTATTTTTCAGTTTCAAGAAGATGAATACATTTTGTATAA 1

RESULT 15
AI333944/c

LOCUS
DEFINITION

AI333944 525 bp mRNA EST 13-FEB-1999
gg27b09.x1 Soares_NhHMPu_S1 Homo sapiens cDNA clone IMAGE:193721
3', mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE

AI333944
AI333944.1 GI:4070503
EST.
human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE
AUTHORS
TITLE

1 (bases 1 to 525)
NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)

JOURNAL
COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgaps-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 990 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 370.

FEATURES
source

Location/Qualifiers
1. 525
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:193721"
/clone_lib="Soares_NhHMPu_S1"
/tissue_type="Pooled human melanocyte, fetal heart, and
pregnant uterus"
/lab_host="DH10B"
/note="Organ: mixed (see below); Vector: pT7T3D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not I;
Site_2: Eco RI; Equal amounts of plasmid DNA from three
normalized libraries (melanocyte 2NBHM, pregnant uterus
NBHPU, and fetal heart NBH19W) were mixed, and ss circles
were made in vitro. Following HAP purification, this DNA
was used as tracer in a subtractive hybridization
reaction. The driver was PCR-amplified cDNAs from pools of
5,000 clones made from the same 3 libraries. The pools
consisted of I.M.A.G.E. clones 260232-265223,
340488-345479, and 484488-489479."

BASE COUNT
ORIGIN

120 a 111 c 171 g 123 t
Query Match 5.0%; Score 461.8; DB 19; Length 525;
Best Local Similarity 97.0%; Pred. No. 3.8e-33;
Matches 513; Conservative 0; Mismatches 12; Indels 4; Gaps 4;

Qy 8346 CTAGAAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCAGG 8405
Db 525 CTAGAAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAACGTTACTCAGG 466
Qy 8406 GCTCTCTCCAGCTTCTAGGAGCTGAGGTTTACCCCTTAAACCCAGGAGCAGCTCCCA 8465
Db 465 GGTCTCTCCAGCTTCTAGGAGCTGAGGTTTACCCCTTAAACCCAA-GGAGCAGCTCCCA 407
Qy 8466 CCTCAGCCCGGGAGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACTTCAGA 8525
Db 406 CCTCAGCCCGGGAGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACTTCAGA 348
Qy 8526 ATTCATATCTTGGAAATCCCAAGCTCCCTGCCCCAAATAAATTCAGTCTGCTTCCAGAA 8585
Db 347 ATTCATATCTTGGAAATCCCAAGCTCCCTGCCCCAAATAAATTCAGTCTGCTTCCAGAA 289
Qy 8586 TTTGGAAATCTAGTTTCTCTCTCTGTTATCCGAGTCTGGGACACAAACTCCGCCCC 8645
Db 288 TTTGGAAATCTAGTTTCTCTCTCTGTTATCCGAGTCTGGGACACAAACTCCGCCCC 229
Qy 8646 CAGCCTATGAGCATCTGAGCCCGCCCTCTTCTTCCGAGAACTGGCCCGGATCAGAGC 8705

```
|||||
Db 228 CAGCCTATGAGCATCTGAGCCCGCCCTCTCTGACGAAACTGGCCCGGATCAGAGC 169
QY 8706 AGGACCTCCCTTCGGACCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCGGAGCATCC 8765
Db 168 AGGACCTCCCTTCGGACCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCGGAGCATCC 109
QY 8766 CGGAGGAAATCTGCAGAGGGGTTAGGAGTGGGTGACAAGAGCCCTGATCTCTCTGTTTT 8825
Db 108 CGGAGGAAATCTGCAGA-GGGTTAGGAGTGGGTGACAAGAGCCCTGATCTCTCTGTTTT 50
QY 8826 GTACATAGATTTATTTTTCAGTTTCCAGAAAGATGAATACATTTTGTTA 8874
Db 49 GTACATAGATTTATTTTTCAGTTTCCAGAAAGATGAATACATTTTGTTA 1
```

Search completed: October 14, 2001, 03:46:30
Job time: 28096 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 08:17:47 ; Search time 2206.06 Seconds
(without alignments)
9952.996 Million cell updates/sec

Title: US-09-380-337-3
Perfect score: 9180
Sequence: 1 CTGCTGTGAACCTCTGGCC.....AGCCACGGCGCGCGCCCG 9180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2230464 seqs, 119590913 residues

Total number of hits satisfying chosen parameters: 4460928

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_New.*
1: /cgn2_6/ptodata/2/pna/PCT_NEW_COMB.seq.*
2: /cgn2_6/ptodata/2/pna/US06_NEW_COMB.seq.*
3: /cgn2_6/ptodata/2/pna/US07_NEW_COMB.seq.*
4: /cgn2_6/ptodata/2/pna/US08_NEW_COMB.seq.*
5: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
6: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
7: /cgn2_6/ptodata/2/pna/US09_NEW_COMB.seq.*
8: /cgn2_6/ptodata/2/pna/US60_NEW_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query %	Match Length	ID	Description
1	9179	100.0	9180	5	US-09-380-337-3
2	9179	100.0	9180	7	US-09-880-107-3461
3	9040.4	98.5	11190	5	US-09-948-941-529
4	9040.4	98.5	11205	5	US-09-948-941-649
5	8067.2	87.9	8185	6	US-09-760-475-3945
6	1425.2	15.5	2770	1	PCT-US01-14827-388
7	1365.2	14.9	2121	1	PCT-US01-14827-5297
8	1312.2	14.3	2772	5	US-09-948-941-5
9	1312.2	14.3	2772	5	US-09-380-337-1
10	1256.4	13.7	2767	5	US-09-948-941-125
11	1252.8	13.6	2168	6	US-09-758-466-258
12	1231.4	13.4	3534	8	US-60-278-561-1743
13	698.6	7.6	3534	8	US-60-278-561-1743
14	639	7.0	639	7	US-09-864-761-22111
15	600.6	6.5	601	5	US-09-948-941-1028
16	600.6	6.5	601	5	US-09-948-941-6056
17	589.6	6.4	601	5	US-09-948-941-1037
18	589.6	6.4	601	5	US-09-948-941-6055
19	540.6	5.9	11557	1	PCT-US01-08656-3267
20	495	5.4	15957	6	US-09-891-498-3
21	473.4	5.2	519	6	US-09-878-178-1236
22	452	4.9	464	7	US-09-864-761-5337
23	414	4.5	15957	6	US-09-891-498-3
24	412.2	4.5	4227	1	PCT-US01-14827-6273
25	400.6	4.4	599	8	US-60-253-651-22088

ALIGNMENTS

RESULT 1

US-09-380-337-3
; Sequence 3, Application US/09380337
; GENERAL INFORMATION:
; APPLICANT: Chandrasekharappa, Settara C.
; Guru, Siradanahalli C.
; Manickam, Pachaiappan
; Collins, Francis S.
; Emmert-Buck, Michael R.
; Debelenko, Larisa V.
; Lubensky, Irina A.
; Liotta, Lance A.
; Agarwal, Sunita K.
; Spiegel, Allen M.
; TITLE OF INVENTION: MEN1, the Gene Associated With Multiple Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses Thereof
; NUMBER OF SEQUENCES: 28
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Townsend and Townsend and Crew LLP
; STREET: Two Embarcadero Center, Eighth Floor
; CITY: San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94111-3834
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/380,337
; FILING DATE: 09-Mar-2000
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/040,269
; FILING DATE: 05-MAR-1997
; APPLICATION NUMBER: WO PCT/US98/04258
; FILING DATE: 04-MAR-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Lockyer, Jean M.
; REGISTRATION NUMBER: 44,879
; REFERENCE/DOCKET NUMBER: 015280-31510005
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (415) 576-0200
; TELEFAX: (415) 576-0300
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:

Sequence 22376, A
Sequence 892, App
Sequence 8463, App
Sequence 216, App
Sequence 6906, App
Sequence 442, App
Sequence 564, App
Sequence 505, App
Sequence 8775, App
Sequence 8703, App
Sequence 34623, A
Sequence 696, App
Sequence 2565, App
Sequence 39315, A
Sequence 695, App
Sequence 607, App
Sequence 2104, App
Sequence 530, App
Sequence 475, App
Sequence 5725, App

LENGTH: 9180 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:

NAME/KEY: -
LOCATION: 1..9180
OTHER INFORMATION: /note= "genomic sequence for MEN1 gene"

FEATURE: NAME/KEY: exon

LOCATION: 1680..1766

FEATURE: NAME/KEY: intron

LOCATION: 1767..2264

FEATURE: NAME/KEY: exon

LOCATION: 2265..2732

FEATURE: NAME/KEY: intron

LOCATION: 2733..4296

FEATURE: NAME/KEY: exon

LOCATION: 4297..4505

FEATURE: NAME/KEY: intron

LOCATION: 4506..4715

FEATURE: NAME/KEY: exon

LOCATION: 4716..4844

FEATURE: NAME/KEY: intron

LOCATION: 4845..5176

FEATURE: NAME/KEY: exon

LOCATION: 5177..5217

FEATURE: NAME/KEY: intron

LOCATION: 5218..5297

FEATURE: NAME/KEY: exon

LOCATION: 5298..5385

FEATURE: NAME/KEY: intron

LOCATION: 5386..6024

FEATURE: NAME/KEY: exon

LOCATION: 6025..6161

FEATURE: NAME/KEY: intron

LOCATION: 6162..6622

FEATURE: NAME/KEY: exon

LOCATION: 6623..6758

FEATURE: NAME/KEY: intron

LOCATION: 6759..7195

FEATURE: NAME/KEY: exon

LOCATION: 7196..7360

FEATURE: NAME/KEY: intron

LOCATION: 7361..7577

FEATURE: NAME/KEY: exon

LOCATION: 7578..8881

SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-380-337-3

Query Match 100.0%; Score 9179; DB 5; Length 9180;

Best Local Similarity 100.0%; Pred. No. 1e-171;

Matches 9180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 CTGCTTTGAACCTCTGGCCTCAAGCAATCCTCCTGCTTCAGCTTCCCAAAGTGTGTAA 60
1 CTGCTTTGAACCTCTGGCCTCAAGCAATCCTCCTGCTTCAGCTTCCCAAAGTGTGTAA 60
61 TTACAGGCATGAGCCTGGCATGAACCTTGACACTATTGAGATATATCTGTCAGGTATTTTG 120
61 TTACAGGCATGAGCCTGGCATGAACCTTGACACTATTGAGATATATCTGTCAGGTATTTTG 120
121 TGGAAATCCTCAACCTGTTTTGCCAGATGTTTTCTCATGATTAGAGGAGATTATAA 180
121 TGGAAATGTCCTCAACCTGTTTTGCCAGATGTTTTCTCATGATTAGAGGAGATTATAA 180
181 ATTTTGGAGAAAATCCAGAGAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 240
181 ATTTTGGAGAAAATCCAGAGAGGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 240
241 ATTTTACTGCTATATACCGAGTGTCTGGAACCTGGCCCATGTAAGTACCAAAAATCTGTT 300
241 ATTTTACTGCTATATACCGAGTGTCTGGAACCTGGCCCATGTAAGTACCAAAAATCTGTT 300
301 TTTTGTGAATGAATAAGCAATAATGAGTGACCGTGAAGTGAAGTGAAGTGAAGTGAAGT 360
301 TTTTGTGAATGAATAAGCAATAATGAGTGACCGTGAAGTGAAGTGAAGTGAAGTGAAGT 360
361 TTCAAAGCGTTGTTGATACAGGCCAGGCACAGTGGCTCACACCTGTATCCAGCAGCTTT 420
361 TTCAAAGCGTTGTTGATACAGGCCAGGCACAGTGGCTCACACCTGTATCCAGCAGCTTT 420
421 TGGAGCCGAGGTAGGAGGATCACTTGAGGTGAGGATTCGAGACCAGCCTGACCAACAT 480
421 TGGAGCCGAGGTAGGAGGATCACTTGAGGTGAGGATTCGAGACCAGCCTGACCAACAT 480
481 GGTGACACCCCTGCTCTACTAAGTAAATAACAAAATAGCCAAAGTGTGTCGAGGCA 540
481 GGTGACACCCCTGCTCTACTAAGTAAATAACAAAATAGCCAAAGTGTGTCGAGGCA 540
541 CCTGTAATCCCGGCTACTTGGGAAGCTGAGGCAGAAGTACACTTGAACCTGGAGGCGAG 600
541 CCTGTAATCCCGGCTACTTGGGAAGCTGAGGCAGAAGTACACTTGAACCTGGAGGCGAG 600
601 AGTTGAGTGAAGCCGAGATACCCCACTGCACTCCAGCCTGAGTGAAGCAGGAGAGCTC 660
601 AGTTGAGTGAAGCCGAGATACCCCACTGCACTCCAGCCTGAGTGAAGCAGGAGAGCTC 660
661 TGCTCAAAACAAATAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 720
661 TGCTCAAAACAAATAAACAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 720
721 AATTTTAGCAGCTTTGGGAGGCTGAGCGGCGAGATCACTTGAAGTTAGGGTTCCAGACC 780
721 AATTTTAGCAGCTTTGGGAGGCTGAGCGGCGAGATCACTTGAAGTTAGGGTTCCAGACC 780
781 AGTCTGGCCCAACATGGTGAACCCCATCTCTACTTAAATAACAAAAGTTTCTGGGTGT 840
781 AGTCTGGCCCAACATGGTGAACCCCATCTCTACTTAAATAACAAAAGTTTCTGGGTGT 840
841 GGTGGCGGAGCCCTATAATCCAGCTACTTGGGAGCTTTTTTTTAAAGCGGAATCTCACTC 900
841 GGTGGCGGAGCCCTATAATCCAGCTACTTGGGAGCTTTTTTTTAAAGCGGAATCTCACTC 900
901 TGTTGCCCGAGGCTGGAGTGCAGTGGCAAGATTTGGCTCACTGAAGCTTCCGCTCCAG 960
901 TGTTGCCCGAGGCTGGAGTGCAGTGGCAAGATTTGGCTCACTGAAGCTTCCGCTCCAG 960
961 GTTCAAGGGATTCCCGCGCCTCAGCCTCCCAAGTACCTCCAGTACCTCCGCTCCAG 1020
961 GTTCAAGGGATTCCCGCGCCTCAGCCTCCCAAGTACCTCCAGTACCTCCGCTCCAG 1020
1021 AAAAAAAAAAAAAAAAAAAAAAATATATATATATATATATATATATATATATATATATAT 1080
1021 AAAAAAAAAAAAAAAAAAAAAAATATATATATATATATATATATATATATATATATATAT 1080

Db 3241 TGAGTAGAGTACAGGATGACCGTGGTGGTAAATTTTGTATTTTTTTTGG 3300
Qy 3301 TAGAGATGGTGTCTACATATGTGGCCAGACTGGTCTTGAACCTCTTGGGCTCAAGTGATC 3360
Db 3301 TAGAGATGGTGTCTACATATGTGGCCAGACTGGTCTTGAACCTCTTGGGCTCAAGTGATC 3360
Qy 3361 TGCCCGCCTCAGTCTCCCAAAATGTGGGATACAGGTGTAGCCACCCCAACTGGTGCC 3420
Db 3361 TGCCCGCCTCAGTCTCCCAAAATGTGGGATACAGGTGTAGCCACCCCAACTGGTGCC 3420
Qy 3421 TATGAAAAATTTTTTTTTTTCAGACGGCGTCTACATCTGTGGCCAGGCTGGAGTGCA 3480
Db 3421 TATGAAAAATTTTTTTTTTTCAGACGGCGTCTACATCTGTGGCCAGGCTGGAGTGCA 3480
Qy 3481 GTGGTGCAATCTGGGCTCAGTGCAGCTCTGCCTCTCTCTTTCATGCCATCTCTCTGGCT 3540
Db 3481 GTGGTGCAATCTGGGCTCAGTGCAGCTCTGCCTCTCTCTTTCATGCCATCTCTCTGGCT 3540
Qy 3541 CCTGGCTCAGCCTCCTGTAGTGTGGGACTACAGAGGCTGCCACCATGCCCTGGCTAATT 3600
Db 3541 CCTGGCTCAGCCTCCTGTAGTGTGGGACTACAGAGGCTGCCACCATGCCCTGGCTAATT 3600
Qy 3601 TTTTGTGGATTTTGTAGAGACGAGGTTTACCATGTTAGCCAGGATGCTCGATCT 3660
Db 3601 TTTTGTGGATTTTGTAGAGACGAGGTTTACCATGTTAGCCAGGATGCTCGATCT 3660
Qy 3661 CTTGACCTCTGATCCGCCCGCTTGGCTGCCAAAGTCTGGGATACAGCGGTGAGCC 3720
Db 3661 CTTGACCTCTGATCCGCCCGCTTGGCTGCCAAAGTCTGGGATACAGCGGTGAGCC 3720
Qy 3721 ACCGCACTCTGTCAAAAATTTGTAGACAGAGAGGGGCTTGACCTCAAAAAGGCTTAAGA 3780
Db 3721 ACCGCACTCTGTCAAAAATTTGTAGACAGAGAGGGGCTTGACCTCAAAAAGGCTTAAGA 3780
Qy 3781 GTCAGGGCTTCCAAAGAGCTTTGCACCAAGCCGGTGTACTGGCAATCCCATCTGGTGT 3840
Db 3781 GTCAGGGCTTCCAAAGAGCTTTGCACCAAGCCGGTGTACTGGCAATCCCATCTGGTGT 3840
Qy 3841 GCCATATTTGAGAAGGAATCAGAGGCTGCTCTCAGCTTTAGCAGGAAAGAGTGCAGAGAT 3900
Db 3841 GCCATATTTGAGAAGGAATCAGAGGCTGCTCTCAGCTTTAGCAGGAAAGAGTGCAGAGAT 3900
Qy 3901 AAATGAGGGTATTTGTGGTGGGTATAGCCAGAGAGTGTGGCCAGCGTCTGTTTTT 3960
Db 3901 AAATGAGGGTATTTGTGGTGGGTATAGCCAGAGAGTGTGGCCAGCGTCTGTTTTT 3960
Qy 3961 TGCCATTCCTGTTTAACTTAGTAGTCAAGTCAAAATGGAATCCCTAAATCCATAGAATA 4020
Db 3961 TGCCATTCCTGTTTAACTTAGTAGTCAAGTCAAAATGGAATCCCTAAATCCATAGAATA 4020
Qy 4021 TATAATAGAGTTGCAGAGAAAGACGAGGTAGGGCCAAAGGCTGGGTCAAGTACAGGATAT 4080
Db 4021 TATAATAGAGTTGCAGAGAAAGACGAGGTAGGGCCAAAGGCTGGGTCAAGTACAGGATAT 4080
Qy 4081 CCAGAAAGGTATCTTTGTGGACATAGAGGGTGTAAACAGGAGAGAGTCTTTGAACAGT 4140
Db 4081 CCAGAAAGGTATCTTTGTGGACATAGAGGGTGTAAACAGGAGAGAGTCTTTGAACAGT 4140
Qy 4141 GGGAGGAAGGATGGAGGATAGTGGCAGGAGAAATCTGAGTTGGGTGCACAGGCTTGG 4200
Db 4141 GGGAGGAAGGATGGAGGATAGTGGCAGGAGAAATCTGAGTTGGGTGCACAGGCTTGG 4200
Qy 4201 AAAGGAGTGGAGGAGTGTGGCCCATCACTACCTGGCCCTTTCCCCCATGTTAAAGCA 4260
Db 4201 AAAGGAGTGGAGGAGTGTGGCCCATCACTACCTGGCCCTTTCCCCCATGTTAAAGCA 4260
Qy 4261 CAGAGGACCTCTTTTCATTAATCTCCCTTCCACAGGACCAAAATTTGACAGCTCCGGTG 4320
Db 4261 CAGAGGACCTCTTTTCATTAATCTCCCTTCCACAGGACCAAAATTTGACAGCTCCGGTG 4320
Qy 4321 TGGCCTTTGCTGTGGTGTGGGCTGCCAGGCGCTCGGGTCTCCGGATGTCACCTCGGCC 4380
Db 4321 TGGCCTTTGCTGTGGTGTGGGCTGCCAGGCGCTCGGGTCTCGGGATGTCACCTCGGCC 4380

Qy 4381 TGTCTGAGGATCATGCTGGTGTAGTGTGGCCCAATTTGGCCCAATTTGGGAGCAGATTTCA 4440
Db 4381 TGTCTGAGGATCATGCTGGTGTAGTGTGGCCCAATTTGGCCCAATTTGGGAGCAGATTTCA 4440
Qy 4441 CCTGGCAGCGCAAGGGCAACGAGGACCCGAGGGCCAGACAGTCAATGCCGCTTTGGGTG 4500
Db 4441 CCTGGCAGCGCAAGGGCAACGAGGACCCGAGGGCCAGACAGTCAATGCCGCTTTGGGTG 4500
Qy 4501 AGCGGATTTGTTCCCTCCCGCCAGCCTTGTCCCTTTCATCTGTAGTAGCCAGCCAC 4560
Db 4501 AGCGGATTTGTTCCCTCCCGCCAGCCTTGTCCCTTTCATCTGTAGTAGCCAGCCAC 4560
Qy 4561 CCAAGGAGCTCCATTTTCTTGGGCAACCCCTTCTTCCCATCACCCACACATAGGAA 4620
Db 4561 CCAAGGAGCTCCATTTTCTTGGGCAACCCCTTCTTCCCATCACCCACACATAGGAA 4620
Qy 4621 GGGAGACAGAGAGCCCTTTTCTGGCTGTCTTCCCTGAAGCAGCAGAGGTGGGC 4680
Db 4621 GGGAGACAGAGAGCCCTTTTCTGGCTGTCTTCCCTGAAGCAGCAGAGGTGGGC 4680
Qy 4681 CATCATGAGACATATGATCTATCCCGCCCTTAAGAGCTGGCTGTACCTGAAGGATCAT 4740
Db 4681 CATCATGAGACATATGATCTATCCCGCCCTTAAGAGCTGGCTGTACCTGAAGGATCAT 4740
Qy 4741 ACATCCGCTGTGACCGCAAGATGGAGGTGGCTTTCATGCTGTGCCATCAACCTTTCCA 4800
Db 4741 ACATCCGCTGTGACCGCAAGATGGAGGTGGCTTTCATGCTGTGCCATCAACCTTTCCA 4800
Qy 4801 TTGACCTGCACACCGACTCGCTGGAGCTTTCAGCTGCAGCAGTGCAGAGGTGAGGCTGAGCCAA 4860
Db 4801 TTGACCTGCACACCGACTCGCTGGAGCTTTCAGCTGCAGCAGTGCAGAGGTGAGGCTGAGCCAA 4860
Qy 4861 TGGGACAGACTGGGCTAGGCAAGCTTGTCTGTGGAGCCCTGGGAGGGGACCTT 4920
Db 4861 TGGGACAGACTGGGCTAGGCAAGCTTGTCTGTGGAGCCCTGGGAGGGGACCTT 4920
Qy 4921 TCCCTTCTGAGCTTTCAGCTTCCCTTCCGAAATGGGTAGTAAATTCCTTGGCTGGC 4980
Db 4921 TCCCTTCTGAGCTTTCAGCTTCCCTTCCGAAATGGGTAGTAAATTCCTTGGCTGGC 4980
Qy 4981 CTTTCCAGGAGCTCTTGGGAGAGTAGAATTTGAGATGTGAAATTTGCTTTCATTTAA 5040
Db 4981 CTTTCCAGGAGCTCTTGGGAGAGTAGAATTTGAGATGTGAAATTTGCTTTCATTTAA 5040
Qy 5041 GGGCTGGTCCAGAAATTTTGGCCCTTCCACATGGTGGTGGTCCCTTGTGGTTCGACCC 5100
Db 5041 GGGCTGGTCCAGAAATTTTGGCCCTTCCACATGGTGGTGGTCCCTTGTGGTTCGACCC 5100
Qy 5101 CCACCTCTGCCGATAGGCTAAGGACCCGTTCTCCTCCTGTTCCTGGCTCATAACTCT 5160
Db 5101 CCACCTCTGCCGATAGGCTAAGGACCCGTTCTCCTCCTGTTCCTGGCTCATAACTCT 5160
Qy 5161 CTCCTTGGCTCTTAGAGCTTCTCTGGCTGTCTATCACCTGGGACATCTTGGAAAGTCT 5220
Db 5161 CTCCTTGGCTCTTAGAGCTTCTCTGGCTGTCTATCACCTGGGACATCTTGGAAAGTCT 5220
Qy 5221 AGTAGAGGAAAGTGGCAGGCTGGCCCTTGGTGGGCGGGGGCTGGGTGGCAGCCTCAA 5280
Db 5221 AGTAGAGGAAAGTGGCAGGCTGGCCCTTGGTGGGCGGGGGCTGGGTGGCAGCCTCAA 5280
Qy 5281 TTATGATCTCTTTCCTAGTACCCCATGAGGAACTTGGGAACTTGGGAACTTGGGAACTG 5340
Db 5281 TTATGATCTCTTTCCTAGTACCCCATGAGGAACTTGGGAACTTGGGAACTTGGGAACTG 5340
Qy 5341 GAGCCACCTCTGGCCGCGCAGACCCATCAGCTTACCACAGGCTGGGCGCATCTAAG 5400
Db 5341 GAGCCACCTCTGGCCGCGCAGACCCATCAGCTTACCACAGGCTGGGCGCATCTAAG 5400
Qy 5401 GAGGTCAGAGGAGAGACCCCTAACAGTGGCTGAGGAGGGGCGCTCATCTGGGCGAGATG 5460
Db 5401 GAGGTCAGAGGAGAGACCCCTAACAGTGGCTGAGGAGGGGCGCTCATCTGGGCGAGATG 5460

Qy	5461	AGAAGAGAACTTTGTGTGTGGGGGGTATCGCCCATCAAGTCTCACATTTGTGTCAACTGT	5520
Db	5461	AGAAGAGAACTTTGTGTGTGGGGGGTATCGCCCATCAAGTCTCACATTTGTGTCAACTGT	5520
Qy	5521	GTGCAGAACTAGTTCACTCAGGGCTCTGTAGGGGTGTCACAGGTTCCCCAGGCTGGGAG	5580
Db	5521	GTGCAGAACTAGTTCACTCAGGGCTCTGTAGGGGTGTCACAGGTTCCCCAGGCTGGGAG	5580
Qy	5581	TGGCAGGGGTGCATTTGTCCCTCAGGCCCTGCCTTTTCTGCCACTGCTTACTGTCTCTTC	5640
Db	5581	TGGCAGGGGTGCATTTGTCCCTCAGGCCCTGCCTTTTCTGCCACTGCTTACTGTCTCTTC	5640
Qy	5641	CTGGAGTATAACAGAGGTCAAATGTGGTAGGACACTGAAGAGGGGTGTCTACTTGGTG	5700
Db	5641	CTGGAGTATAACAGAGGTCAAATGTGGTAGGACACTGAAGAGGGGTGTCTACTTGGTG	5700
Qy	5701	GGTGTAGGTGGGAGGAGGGCCATTGGGCTGGGCTTGAAGTCTTTCGTGATGTGTAGAA	5760
Db	5701	GGTGTAGGTGGGAGGAGGGCCATTGGGCTGGGCTTGAAGTCTTTCGTGATGTGTAGAA	5760
Qy	5761	GAGTGTCTGAGAAAGAAAGGCCCTGAGCTCGGAGGCGCCACCCCTGCAGTCTG	5820
Db	5761	GAGTGTCTGAGAAAGAAAGGCCCTGAGCTCGGAGGCGCCACCCCTGCAGTCTG	5820
Qy	5821	CCCCAGGCTCAGCCAGCAGTCTCTGTAGACCCAGGAGGAGACAGGTAGAAGGCTGGC	5880
Db	5821	CCCCAGGCTCAGCCAGCAGTCTCTGTAGACCCAGGAGGAGACAGGTAGAAGGCTGGC	5880
Qy	5881	AGCGAGTGAGGTGGGAGTGGAGATGGAGAGACTCCCTGGGATCTTTCCTGTGCCCTT	5940
Db	5881	AGCGAGTGAGGTGGGAGTGGAGATGGAGAGACTCCCTGGGATCTTTCCTGTGCCCTT	5940
Qy	5941	CTGGGTGTGCCCTGTGTGGGGCAATTGTGCCAGAGGGCAGCTGGGGCTGCTCCCTGAGG	6000
Db	5941	CTGGGTGTGCCCTGTGTGGGGCAATTGTGCCAGAGGGCAGCTGGGGCTGCTCCCTGAGG	6000
Qy	6001	ATCCTCTGCCCTACCTTCCATCCAGGGCATGTGCCTCAGCCAAAGACTACTATCGGGATGAA	6060
Db	6001	ATCCTCTGCCCTACCTTCCATCCAGGGCATGTGCCTCAGCCAAAGACTACTATCGGGATGAA	6060
Qy	6061	CACATCTACCCCTACATGTACCTGGCTGGCTACCACTGTCCCAACCGCAATGTCCGGGAA	6120
Db	6061	CACATCTACCCCTACATGTACCTGGCTGGCTACCACTGTCCCAACCGCAATGTCCGGGAA	6120
Qy	6121	GCCCTCAGGCCCTGGCGGACACGGCCACTGTCTCAGGAGTGAGGATCCCCCTACTAG	6180
Db	6121	GCCCTCAGGCCCTGGCGGACACGGCCACTGTCTCAGGAGTGAGGATCCCCCTACTAG	6180
Qy	6181	GGCCTGCAGGCTGTCTTTCTTTCCCTCCATCAGTTTCCAAACCCCTCGTCCAGGACTG	6240
Db	6181	GGCCTGCAGGCTGTCTTTCTTTCCCTCCATCAGTTTCCAAACCCCTCGTCCAGGACTG	6240
Qy	6241	AGGCCTGGCTCCACGCCCCCATCCCTTTTCCATCCAGTCCCTAGCAGCAAGGCCACCAT	6300
Db	6241	AGGCCTGGCTCCACGCCCCCATCCCTTTTCCATCCAGTCCCTAGCAGCAAGGCCACCAT	6300
Qy	6301	TACCCAGGAGTAGGACCCCTGATTAAAGGTGTACATCTTTCCCTCCCTCCCTCTCTCTC	6360
Db	6301	TACCCAGGAGTAGGACCCCTGATTAAAGGTGTACATCTTTCCCTCCCTCCCTCTCTCTC	6360
Qy	6361	CTAATTTTTTTTTTCTCAGAACAGTCTCAAAATCTCCAATGTTTAAACCACCATCATCCAGC	6420
Db	6361	CTAATTTTTTTTTTCTCAGAACAGTCTCAAAATCTCCAATGTTTAAACCACCATCATCCAGC	6420
Qy	6421	AGTGGGACTTCCACCTCGGCCCCATGCCCTCCTCATTTCTTGTCTTTCTTCTCTCTGGG	6480
Db	6421	AGTGGGACTTCCACCTCGGCCCCATGCCCTCCTCATTTCTTGTCTTTCTTCTCTCTGGG	6480
Qy	6481	CTGACCCAGACAGCATCATTTTGCAGTGAGGACCCCACTACTCCCCCAGCCCTGGGG	6540
Db	6481	CTGACCCAGACAGCATCATTTTGCAGTGAGGACCCCACTACTCCCCCAGCCCTGGGG	6540
Qy	6541	CTCCATCCCCCGCCAGGTCCTCTGGGCTACCCCGCATGGTGAGACCCCTTTCAGACCCCTAC	6600

Db	6541	 CTCCATCCCCCGCAGGTCCCTTGGGCGTACCCCAGTGSTGAGACCCTTTACAGCCCTAC 	6600
Qy	6601	ACAGACCCCACACTGCTCTCACAGCTACAACACTACTTCGCCGGGAAGACGAGGAGATCTACAAGG 	6660
Db	6601	AGAGACCCCACACTGCTCTCACAGCTACAACACTACTTCGCCGGGAAGACGAGGAGATCTACAAGG 	6660
Qy	6661	AGTTCTTTGAAGTAGCCAATGATGTCATCCCCAACCTGCTGAAGGAGGCAGCCAGCTTGC 	6720
Db	6661	AGTTCTTTGAAGTAGCCAATGATGTCATCCCCAACCTGCTGAAGGAGGCAGCCAGCTTGC 	6720
Qy	6721	TGGAGCGGGCGAGAGCGCGCGGGGAGCAAAAGCCAGGTGAAGGCTGGAGCTCAGCC 	6780
Db	6721	TGGAGCGGGCGAGAGCGCGCGGGGAGCAAAAGCCAGGTGAAGGCTGGAGCTCAGCC 	6780
Qy	6781	TGTTGTCCAGCCTCCCAACCTGGACAGGGCTCCCTTCCACAGGGCCATTGGGGCTGCATGTA 	6840
Db	6781	TGTTGTCCAGCCTCCCAACCTGGACAGGGCTCCCTTCCACAGGGCCATTGGGGCTGCATGTA 	6840
Qy	6841	CGGGATTAGGGATGGCAGGAGGAAGGTGGCCCTGAGCAGACAGCTATGTTTTCCCTTTTGCCT 	6900
Db	6841	CGGGATTAGGGATGGCAGGAGGAAGGTGGCCCTGAGCAGACAGCTATGTTTTCCCTTTTGCCT 	6900
Qy	6901	ATAACTGAGGTCCCTGGGCCACGTTGGACGGGACTGAAGGTATTTTAGAGGTTTCTACCC 	6960
Db	6901	ATAACTGAGGTCCCTGGGCCACGTTGGACGGGACTGAAGGTATTTTAGAGGTTTCTACCC 	6960
Qy	6961	TGTTGCTTTCAGTTTCATGSCCAGACTCCCTCCCTCAGCTGAGGGGTGGAGTAGGGATGG 	7020
Db	6961	TGTTGCTTTCAGTTTCATGSCCAGACTCCCTCCCTCAGCTGAGGGGTGGAGTAGGGATGG 	7020
Qy	7021	TACGTCTGGCTATGATTTGGCTTTTATAAAGGAAAAGAGGTTCTAAGAATGTTCCCAACC 	7080
Db	7021	TACGTCTGGCTATGATTTGGCTTTTATAAAGGAAAAGAGGTTCTAAGAATGTTCCCAACC 	7080
Qy	7081	TATGCTTTACCTTTCTGAGGACAGGGGTCTTTCGCTTAGGTGGGGGCGCTGGCCTGTGCC 	7140
Db	7081	TATGCTTTACCTTTCTGAGGACAGGGGTCTTTCGCTTAGGTGGGGGCGCTGGCCTGTGCC 	7140
Qy	7141	TCCTGCTAAGGGGTGACTAAGAGACTCATCTGTGCCCTCCCTTCCCTCCGTCGACAGGSCAC 	7200
Db	7141	TCCTGCTAAGGGGTGACTAAGAGACTCATCTGTGCCCTCCCTTCCCTCCGTCGACAGGSCAC 	7200
Qy	7201	CCAGAGCCAAAGGTTCCGCCCTCCAGGACCCCTGAGTGCCTGCCACACCTGCTGGGATTTCTA 	7260
Db	7201	CCAGAGCCAAAGGTTCCGCCCTCCAGGACCCCTGAGTGCCTGCCACACCTGCTGGGATTTCTA 	7260
Qy	7261	CGACGGCATCTGCAATGGAGGAGGGCAGTCCACCGCTGTGCTGCACGTGGGCTGGGC 	7320
Db	7261	CGACGGCATCTGCAATGGAGGAGGGCAGTCCACCGCTGTGCTGCACGTGGGCTGGGC 	7320
Qy	7321	CACCTTTCCTTGTGCAGTCCCTTAGCCGCTTTTGGGGACAGGTGAGGACAGCTGCCACAGA 	7380
Db	7321	CACCTTTCCTTGTGCAGTCCCTTAGCCGCTTTTGGGGACAGGTGAGGACAGCTGCCACAGA 	7380
Qy	7381	GGTCTGGGCACCTACAGTGGTGTGACAGCAGCCAGGGCTTCTCAGACTTTTCTGGCCCCAGG 	7440
Db	7381	GGTCTGGGCACCTACAGTGGTGTGACAGCAGCCAGGGCTTCTCAGACTTTTCTGGCCCCAGG 	7440
Qy	7441	GGAGCATCTGCCCATTCGCCCTTCGGTGGCATGGGACTGAGACCCCTGGGTGGGATGGG 	7500
Db	7441	GGAGCATCTGCCCATTCGCCCTTCGGTGGCATGGGACTGAGACCCCTGGGTGGGATGGG 	7500
Qy	7501	ATNGGCCAGAGCAGGGTCCCTGGAGTTCACGCCACTGGCCGCAACCTTGCTCTCACCTTGC 	7560
Db	7501	ATNGGCCAGAGCAGGGTCCCTGGAGTTCACGCCACTGGCCGCAACCTTGCTCTCACCTTGC 	7560
Qy	7561	TCTCCCACTGGCCCAAGGTGCGCATAGTGAGCCGAGAGGCCAGAGCCGAGGCGG 	7620
Db	7561	TCTCCCACTGGCCCAAGGTGCGCATAGTGAGCCGAGAGGCCAGAGGCGG 	7620
Qy	7621	CCGAGGCCGAGGACCGCTGGGGCAGGAAGCCCGGGAAGCCGCGGCGGGGCCACGCG 	7680

Db 7621 CCGAGGCCAGAGCCGTGGGGGAGAGAGCCCGGGAAGCCGGGGGGGGGGCCACAGC 7680
Qy 7681 GGGAGTCCAAAGCCAGAGAGCCCGCCCGCCCAAGAGCCAGCAGCTGGACAGAGGCGCTGG 7740
Db 7681 GGGAGTCCAAAGCCAGAGAGCCCGCCCGCCCAAGAGCCAGCAGCTGGACAGAGGCGCTGG 7740
Qy 7741 GCACGGCCAGGGTGCAGTGTACAGAGACCCCGCCCGGAAAGCCCTCTGGGACTGTGCGTGGCA 7800
Db 7741 GCACGGCCAGGGTGCAGTGTACAGAGACCCCGCCCGGAAAGCCCTCTGGGACTGTGCGTGGCA 7800
Qy 7801 CAGCCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCAGCAGCATCACAC 7860
Db 7801 CAGCCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCAGCAGCATCACAC 7860
Qy 7861 GCGCGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGAGCATGAAGGAGTGC 7920
Db 7861 GCGCGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGAGCATGAAGGAGTGC 7920
Qy 7921 TGGTGGCCACCAAGATCAACTCGAGCGGCATCAAGCTGCAACTCACGGCAGCTGCGAAG 7980
Db 7921 TGGTGGCCACCAAGATCAACTCGAGCGGCATCAAGCTGCAACTCACGGCAGCTGCGAAG 7980
Qy 7981 TGCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTGTACTACTCTGTCTTCTCTCAAGC 8040
Db 7981 TGCAGATGAAGAGCAGAAAGTGTCCACCCCTAGTGTACTACTCTGTCTTCTCTCAAGC 8040
Qy 8041 GGCAGCGAAAGGCTCTGAACTACTGGGAGCTTCGGAGCCGCTGTGGGAGCCAGGCTC 8100
Db 8041 GGCAGCGAAAGGCTCTGAACTACTGGGAGCTTCGGAGCCGCTGTGGGAGCCAGGCTC 8100
Qy 8101 GCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCGCTC 8160
Db 8101 GCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCCCCCAGCCCAAGGGGACAGGCGCTC 8160
Qy 8161 ACCTTACCCAAACCTAGTTCGCGGTCCGAGTACAGTCTGTATCAAAACCCAGCATTT 8220
Db 8161 ACCTTACCCAAACCTAGTTCGCGGTCCGAGTACAGTCTGTATCAAAACCCAGCATTT 8220
Qy 8221 TCTCCAGTCAAGACCCAGGCTCTGCCCCAGTCTGTAGATATAGGTCTCTTCTCCAG 8280
Db 8221 TCTCCAGTCAAGACCCAGGCTCTGCCCCAGTCTGTAGATATAGGTCTCTTCTCCAG 8280
Qy 8281 AATCCAGCCGCGCCAAATGAAACCTCAGCGTGGGTCTTAATTACAGCTTTTAAAGGCC 8340
Db 8281 AATCCAGCCGCGCCAAATGAAACCTCAGCGTGGGTCTTAATTACAGCTTTTAAAGGCC 8340
Qy 8341 AGCCCTTAGAACCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTTAC 8400
Db 8341 AGCCCTTAGAACCAAGCTCTCTCTCGGAACCGCTCAGCTAGAGCCAGACCAAGTTAC 8400
Qy 8401 TCAGGGCTCCTCCAGCTGTAGAGCTGAGGTTTTCACCCCTTAACCCAGGGAGCACAGG 8460
Db 8401 TCAGGGCTCCTCCAGCTGTAGAGCTGAGGTTTTCACCCCTTAACCCAGGGAGCACAGG 8460
Qy 8461 TCCCACTTCCAGCCGCGGAGCCTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACT 8520
Db 8461 TCCCACTTCCAGCCGCGGAGCCTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACT 8520
Qy 8521 TCAGAAATTCATATCTTGGGAATCAAGCTCCCTCGGCCCAATTAAGTTCAGTCTGCTTC 8580
Db 8521 TCAGAAATTCATATCTTGGGAATCAAGCTCCCTCGGCCCAATTAAGTTCAGTCTGCTTC 8580
Qy 8581 CAGAAATTTGGAATCCTAGTTTTCCTCTCTGATCCGAGCTCTGGGACACAAACATCC 8640
Db 8581 CAGAAATTTGGAATCCTAGTTTTCCTCTCTGATCCGAGCTCTGGGACACAAACATCC 8640
Qy 8641 GCGCCAGCCTATGAGCATCTGAGCGCCGCGCTCTTCTGACGAATCTGCGCCCGGATC 8700
Db 8641 GCGCCAGCCTATGAGCATCTGAGCGCCGCGCTCTTCTGACGAATCTGCGCCCGGATC 8700
Qy 8701 AGAGCAGAGACCTTCCCTCGGACCTCTGGGAACCTCCAGAGGTCACGCCATCTCGGAG 8760
Db 8701 AGAGCAGAGACCTTCCCTCGGACCTCTGGGAACCTCCAGAGGTCACGCCATCTCGGAG 8760

RESULT 2

US-09-880-107-3461
; Sequence 3461, Application US/09880107
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3461
; LENGTH: 9180
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. U93237
; NAME/KEY: unsure
; LOCATION: (1)..(9180)
; OTHER INFORMATION: n = a or c or g or t
US-09-880-107-3461

Query Match 100.0%; Score 9179; DB 7; Length 9180;
Best Local Similarity 100.0%; Pred. No. 1e-171;
Matches 9180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CTGGTCTTGAACCTCGGCTCAAGCAATCTCTCCCTTCAGCTTCCCAAGCTGTGTAA 60
Db 1 ctggcttgaaacctcggctccaagcaatctctcccttcagcttcccaagctgtgttaa 60
Qy 61 TTACAGGATGAGCTGGCATGAAGTTCAGCTATTGAGATATCTGCTGAGTATTG 120
Db 61 ttacaggatgagctggcatgaagtctcagctattgagatatactgctgagatttg 120
Qy 121 TGGAAATGTCCTCAACTCTGTCTTCCAGATGTTTCTCATGATTAGAGGAGTTATA 180

Db 121 |||||tggaatgtccctcaactctgttttgcagatgtttctcatattagagagagtataa 180
QY 181 ATTTTGGAGAAATCCAGAGAGTGAAGAGTGAAGTAGGGCAGAAATTTAATCTGTTTT 240
Db 181 |||||atlttgaggaaatccagagagtggaaggtgaagtgagggcagaaatcctcaactgttt 240
QY 241 ATTTACTGCTATATACCGAGTGTCTGGAACTTGGCCCATGGTAAGTACCAAAAATCTGTT 300
Db 241 atttactgctataccgagtgctggaacttggccatgtgtaagtaccaaataatctgtt 300
QY 301 TTTTGTGAATGAATAGCAAAATAATGAGTAGACCGTGGAAATTTAGTATATTTCAAAGT 360
Db 301 ttttttgaaatgaataagcaataataatgagtgaacgtggaaatttagtattatttcaagt 360
QY 361 TTCAAAGCGTTGTGTATACAGGCCAGGCACAGTGGCTCACACCTGTAAATCCCAGCACTTT 420
Db 361 ttcaaagcgtgtgtatatacaggccagggcacagtggctccacacctgttaatcccgacattt 420
QY 421 TGGAGGCCGAGTAGGAGGATCACTTGAAGTCAGGAGTTCGAGACCAAGCTGACCAACAT 480
Db 421 tggagggcaggtaggagagatcacttgaggtcagaggttcgagaccagcctgaccaaat 480
QY 481 GGTGACACCCCTGTCTACTACTAGTAAATACAAAATTAGCCAAAGTGTGGTGCCAGGCA 540
Db 481 ggtgacacccctgtctctactaagtaataatacaaaaaattagccaagtgtgtgtgcaaggca 540
QY 541 CTTGTAATCCCGCTACTTGGGAAGCTGAGGCAGAGAAATCACTTGAACCTGGGAGGACG 600
Db 541 cctgtaatcccgctactctgggaagcttgaggcagagaaatcaacttgaacctggagggcag 600
QY 601 AGGTTGCAGTGAGCCGAGATCACCCCACTGCACCTCCAGCCTGAGTGACAGAGCAGACTC 660
Db 601 aggttgcagtgagccgagatccacccactgcactccagcctgagtgcagagcgagactc 660
QY 661 TGCTCAAAACAATAAACAATAAATACTACTTCTTGGCCGGTGAAGTGTTCAGCCCTGT 720
Db 661 tgtctcaaaaaataaacaataaactactcttggccgggtgaagggtgttcaacgcctgt 720
QY 721 AATTTTAGCACTTTGGGAGCTGAGGCGGCAGATCACTTGAAGTTAGGGGTTTCGAGACC 780
Db 721 aattttagcaactttgggaggtcagggcagagatacacttgaaggttggggttcgagacc 780
QY 781 AGTGTGCCCAACATGTTGAACCCCACTCTCTACTTAAATACAAAAGTTTCTGGGTGT 840
Db 781 agtctgcccaacatggtgaaaccccatctctacttaataatacaaaaagttttctgggtgt 840
QY 841 GGTGGCGGACGCTATATATCCAGCTACTTGGGACTTTTAAAGACGGAATCTACTC 900
Db 841 ggtggcgagcgcctataatccagctacttgggactttttttaagacggaatctcactc 900
QY 901 TGTGTCCCAAGCTGGAGTGAGTGCGCAAGATTCTGGCTCACTGAAGCCTCCGCCCTCCAG 960
Db 901 tgttgcaccaggtcggagtgagtgggcagatctctggtctactgaagctccgcctccag 960
QY 961 GTTCAAGGGATTCCCGCGCCTCAGCCTCCCAAGTAGCTGGGAATCCCTGTCTCTGCAA 1020
Db 961 gttcaagggattccccgcgcctcagcctcccaagtacttgggaatccctgtctctgcaa 1020
QY 1021 AAAAAAATAAAAAAATAAATAATATATATATATATATATATATATATATATATATAT 1080
Db 1021 aaaaaaataaaaaaataaataataataataataataataataataataataataataata 1080
QY 1081 TGTGTGTGTAT 1140
Db 1081 tgtgtgtgttatgta 1140
QY 1141 GAGTGTGGCGCAGCCTGTGATCACAGCTACTCGGAGGCTGAGGCACGAGAATCGCTT 1200
Db 1141 gagt 1200
QY 1201 GAGCCCGTGAAGTGCAGGTGCAGTGAGCCGAGATCGAGCCACTTGCATTCAGCCTGGGC 1260
Db |||||

Db 1201 gagccgtgaagtgcaggtgcagtgagcccgatcgagccactgcaattccagcctgggc 1260
QY 1261 GAAAGAGAAAGACCGTGTCTCAAAACAAACAAACAAAGCTACTCTTAGCACGTGTAGA 1320
Db 1261 gaaagagaaagacggtctctcaaaaaaacaacaaagctactcttagacacgtgttaga 1320
QY 1321 GTATCTCGCGGGGGAAGTGGGAAACGAGTGTGCACACAGAGTAGGCATCTTTATATGT 1380
Db 1321 gtatctcggggcgggaagtggaaacgagtgctgcacacagtagagcatctttatatgt 1380
QY 1381 TAACAGACACTGATACCCAGCTAAAGCGGCTGAACACATTTACTCTCTGCGACGTGTTAA 1440
Db 1381 taacagacactgatataccacagctaaagcggctgaacatctactctctgagtggttaa 1440
QY 1441 AAGTATCTGTTTTCTCATATTGTTTTTAAATTTTCTGGATCAAGCAACCTGATC 1500
Db 1441 aagtatctgttttctcatatattgttttattttaatttttcttgatcaagaacctgac 1500
QY 1501 TTTTTCCTCATATACTTCCGACCGACCCGTCACAGCAAAAACCCGCGAGAGTCTGGCGACC 1560
Db 1501 ttttctctataacttgccgacgcgacccgtgacagcaaaaccggcagagctcgcgacc 1560
QY 1561 TCCCACCCCGAGTCTGAGTAGTGCCTCCCGGACTACATTTCCAGAGGCACTGCGGGC 1620
Db 1561 tcccaccccgagctgcaggtagtgcccccgactacatttccagaaagcacgtcgggc 1620
QY 1621 ACGTTCTCCTGCTGGTGGCTGAAGGAAAGGGCAATCCCTGAGTATCTCGGGAAGGAG 1680
Db 1621 acgttctctcgttcggtcgtgaaggaaagggccaatccctgaatctctcggaagag 1680
QY 1681 GTGTCCGAGCGCGGACCTAGAGATCCAGAACCCACAGCGCAGCGCCCGCCGCCA 1740
Db 1681 gtgcggagcgcgaccttagatcccaagaagccacagcgacgcgcccgcga 1740
QY 1741 CTATTTCCAGCTCTCGGSGCAGGGTGGGCCCCAGACTCCACTTCCGCGGGTAGTGC 1800
Db 1741 ctatttccaggtctgcgggcaggggtggggccagactccacttccgcgggttagtgc 1800
QY 1801 GACCTTAGGGCGGGACTTTCATGTCCAGCAGGCTCCGGGCGGTGCGCGCGGTGCCT 1860
Db 1801 gaccttagggcgggacttcatgtccccagcagctccggcggtgcgcgcggtgtcct 1860
QY 1861 AGTGTGGATGTAAAGCGGAGGTGGGCGGAGCGGAGCCGAGGCACTCTCTTTGGGG 1920
Db 1861 agtgtgggtatgaagcgggaggtgggcgaggggagccagagcactctcttgggg 1920
QY 1921 TTTGGSGGCTTGACTGCTGCGCTTCTGACAGACTTTACAGCCCGCGGGGACAGT 1980
Db 1921 tttgggggttgacctgggtgcgtttcttggacagactttacagccccccggggacagt 1980
QY 1981 CGTAGAGAGGGGGCGGGCGGCCCATTTGGGGCTCTCATTTGGGGTCTTGGGCGCACCCC 2040
Db 1981 cgtagaagggggcgggcgccatggggctcctcattgggtgtccttggggcgaccccc 2040
QY 2041 ATCGGGTACCGGGCGTCCCGGAATTTGGGGGACAAAGGCTCTGAGTCTCGGCTGAG 2100
Db 2041 atcgggtaccggcgctcccggaattgtggggacaaaagctctgcagctcgtcgtag 2100
QY 2101 GGGTCTCACCCACAANAAGAGGGGAAGCCGTGACGAGAGCTGAAGGGTGGGGAAGCA 2160
Db 2101 ggggtctaccgacaaaagaggggaagccggtgagcagaggtgaagaggggtgggaagca 2160
QY 2161 GGGGAGTGTGCTGTGTGGGGGGGTGGAACCTTAGCGGACCTTGGGAGGAGGCTCCC 2220
Db 2161 ggggagctgt 2220
QY 2221 CGGCGGAACCTGCCCGACCTCCCTCCCGCGGTGCTTTCAGGCGCCGCCCCACCCGCC 2280
Db 2221 cgcggaacctgtcccgacctccctccccggcttgccttcagggcgccgccacccgccc 2280
QY 2281 CGCGCCCATGGGCTGAAGGCCGCCAGAAACGCTGTTCCCGCTGCGCTCCATCGAGCA 2340
Db 2281 cgcgcccatggggtgaaggcgcccgagagcgtgttctccgcgtgcgtctccatcgacga 2340

|||||
Db 4501 agcgggtattgttccccccagcctgtcccccttcaactgtagtagcccaagccac 4560
QY CCAAGGGACTCCATTCTTGGCCGACACCCCTTCTTCCCATCCACCACCATAGGAA 4620
Db ccaaggagacccattctcttgggcccacacccctctctcccatcacccacccataggaa 4620
QY GGGAAGACAGAAAGAGCCCTTTTCCCTGGCTGTCATTCCCTGGAAGCAGCAGAGGTGGGC 4680
Db ggaagacagaagagcccccttccctggctcatlccccctgaagcaggcacagggtgggc 4680
QY CATCATGAGACATANTGATCTATCCCCCTTAAGAGCTGGCTGTAACCTGAAGAGTAT 4740
Db catcatgagacataatgatctcatcccccttaagagctggctgtacctgaaaggatcat 4740
QY ACATGGCTGTGACCGCAAGATGGAGTGGCTTCATGTTGTGTGCCATCAACCCCTTCCA 4800
Db acatggctgtgacccgcaagatggaggtgttcatgggtgtgccatcaaccccttcca 4800
QY TTGACCTGCACACGACTCGCTGGAGCTTCTGCAGCTGCGCAGGCTGAGGGCTCAGCCAA 4860
Db ttgacctgcacaccgactcgctgagcttctgcagctgcagcaggtgagggctlgagccaa 4860
QY TGGGGCAGGACTGGCTAGGCCAGACTTGACTTGTCTGTGGGACCCCTGGGGCAGGGCACTT 4920
Db tggggcaggagctggctagccagacttgacttgctgtgggacccctgggcaggggcactt 4920
QY TCCTTTCTCAGCTTCAGCTTCCCTCTCGGAAATAGGTTAGTAATTCCTGGCCCTGGC 4980
Db tcccttcttgagcttcagcttccccctctctggaaaaatgggtctagtaattcctggctggc 4980
QY CTTTCCAGGGCTCTTGGGAGAGTAGAATTTGAGATGTGAATTTGCTTTGACTCATATAA 5040
Db ctttccagggtctlgggagagtagaattgagatgtgaattgcttgaacctattaaa 5040
QY GGCTGTGCCAGAAATTTGGCCCTTCCACATGGTGGTGGTCCCTGTGTTGGTCTGACCC 5100
Db gggctgtccagaaatttggcccttccacatggtgggtggctcctgttgggtctgacc 5100
QY CCACCTCTGCCCATAGGCTAAGGACCGTTCTCCCTCTGTTCCGTTGGCTCATACTCT 5160
Db ccacctctgcccagataggctcaaggacccgttctctctctctcttccgtgggtctaaactct 5160
QY CTCCTTGGCTCTAGAACTGTCTGGCTGCCTATGACTGGGACATCTGGAAGGTC 5220
Db ctctctggctctagaagctgtctggctgtctctatgatcctgggacatctgaaaggctc 5220
QY AGTAGAGGAAGTGGCCAGGCTCGGCTGTGTAGGCGGGGGGGCTGGGTGGCAGCCGTGAA 5280
Db agtagagggaagtggccaggctgcgctggctggtagggccggggggtgggtggcagcctgaa 5280
QY TTATGATCTCTTCTAGGTACCCATGGCCCTTAGGGAACCTGCAGATCTAGAGGAGCTG 5340
Db ttatgacctctctagttaccccaaggccttgaggacccctggcagatctagaggagctg 5340
QY GAGCCACCCCTCGCCGGCCAGACCCACTCACCTCTTACCACAGGTGGGGCATCTAAG 5400
Db gagccacccctggcggccagagcccaactcacctctaccacaagggtggggcatctaa 5400
QY GAGGGTCAGAGGAGGACCCCTAACAGTGGCTGAGGCGGGGGCCCTCATCTGGGCAGATG 5460
Db gaggggtgcagaaggagacccctaacagtggctgaggcaggggccctcatctgggcagatg 5460
QY AGAAGAGAACTTTGTGTGTGGGGGTATCGCCCATCTCAGCTCTCAGTTTCTGTCAACTGT 5520
Db agaagagaaacttgggtgttgggggggtatcgcccataccagctccacttctgtcaactgt 5520
QY GTGCAGAAATCAGTTCAGTCAAGGCTGTCTGAGGGGTGTCCAGGGTTCCCCAGGCTGGGAG 5580
Db gtgcagaatcagttcagtcagtcaggtgtctgaggggtgtccagggttccccagcctggag 5580
QY TGGCAGGGGTGCTATTTGTCCCTCTAGCCCTGCTTTTGTGCCAAGTGTCTTACTGTCTCTC 5640
|||||

Db 5581 tggcaggggctgcatctgttccccctcagccctgctctttttgcccactgcttactgtcttc 5640
QY CTGAGGTATTAACAGAGTCAAAATGTGTAGAGCACTGAAAGAGGGGTGTTCACTTTGGTG 5700
Db ctggagctataacagaggtcaaaagtgtgtaggagcacctgaagagggggtcttactctgggtg 5700
QY GGTCTAGGTGGGAGGAGGCCCATTTGGGCTGGCTTTGAAAGTCTTTGGTGTATGTGTAGAA 5760
Db ggtgtagggtggggagagggccattgggtgggtgagcttgaaagttcttgggtgtagtagaa 5760
QY GAGTGTCTGAGAAAGAGAGAGGGCCCTGAGCTCGAGGGCAGGCCACCCCTCGAGTCTG 5820
Db gagtgtctgagaaagagagggccctgagctcgagggcagggccccccctcagctctg 5820
QY CCCAGGCCCTCAGCAGCAGTCTGTAGACCCAGGAGGAGAGCAGGTAGAACGGCTGGC 5880
Db cccagggcctcagcagcagctcctgttagaccagggagagagaccaggttagaaagggctggc 5880
QY AGCGAGTGGAGGTGGAGTGGAGTGGAGAGTCCCTGGGATCTTCTGTGTGGCCCTTT 5940
Db agcgagtgagggtgggagtgagatggagaggaactcctcctggatcttctctgtggccctt 5940
QY CTGGGTGTGCCCTGGTGGGCATTTGTGCCAGCAGGCAGCTGGGGCTGCTCCCTGAGG 6000
Db ctgggtgtgcccgtgggtggggcatctgtgccagcaggcagctggggctgctccctcaggtg 6000
QY ATCTCTGCTCAGCTACCTCCATCCAGGGCATTTGCCAGCAGGCACTACTATCGGGATGAA 6060
Db atctctgctcactccatccatccaggcatctgctcagcagaagaccactactatcgggatgaa 6060
QY CACATCTACCCCTACATGTATACCTGGCTGGCTTACCACTGTCTCGCAACCCGCAATGTGCGGAA 6120
Db cacatctaccctacatgtacctggctggctacacactgtcgcaaccgcaatgtgctgggaa 6120
QY GCCCTGCAGGCCCTGGGGGACACGGCCACTGTCTATCCAGAGTGCAGGATCCCCCTACTAG 6180
Db gccctgcaggcctgggcggacagccactgtctacaggagtgaggtacccccactag 6180
QY GGCTGCAGGCTGTGCTTTCTTCCCTCCATCATGTTTCCAAACACCCCTGCTCCAGGACTG 6240
Db ggctgcagcctgtctcttcttccctccatcagtttccaaaccacccctgttccaggactg 6240
QY AGCCCTGGCTCCACGCCCCCATCCCCCTTCCATCCAGTCCCTAGGCAGCAGGCCACCAT 6300
Db agccctggctcccaagcaccatcccccttccatccagtcctcaggcagcaaggccacat 6300
QY TACCAGGAGGTAGGACCCCTGATTAGGTGTACATCTTTTCCCTCCTCCTCCTCCTCCTC 6360
Db taccagggaggtaggagaccctgattaaaggtgcacatcttctcctcctccctcctcctc 6360
QY CTAATTTTTTTTTTCTCAGAACAGTCTCAATCTCCAATGTTTAAACCACCATCATCCAGC 6420
Db ctaatttttttttctcagaaacagctccaaatctccaaatgtttaaaccacacatccacagc 6420
QY AGTGGGACTTCCACCCCTCGGCCCCCATCCCCCTTCCATCTTCTTGTGTTTCTTCTCTGGG 6480
Db agtgggacttccaccctcgcccccataccccctcctcattcttcttcttctctctggg 6480
QY CTGACCCAGACAGATCATTTTGCAGTGCAGGACCCCACTACTTCCCCAGCCCTGGGG 6540
Db ctgaccagacagatcatcttgcagtgaggaccccaactactccccccagccccctgggg 6540
QY CTCATCCCCCGCAGTCTCCTGGGGCTACCCCGATGGTGAGACCCCTTTCAGACCCCTAC 6600
Db ctcatcccccccgccagggtccccctggggtacccccgaatgggtgagacccttccagaccctac 6600
QY AGAGACCCCACTGCTCTCACAGCTTACAACCTACTTCCCGGGAAGCAGGAGATCTTACAAG 6660
Db agagaccccaactgctctcagctacaaactactgcccggggaagcagagagatctacaagg 6660
QY AGTTCTTTGAAGTAGCAATGATGTATCCCAACCTGCTGAAGGAGGCAGCCAGCTTGC 6720
Db agttctttgaagtagccaatgatgtcatcccccaacctgctgaaggaggcagccagcttgc 6720

Db 8881 atataaagcgaagtcacatgtttatctgggaaattggggtgggagtggaagcgc 8940
QY 8941 CCCTTCTCCCTTGTCTTCTGGCTCCCGGACTTTCGCTCCCTACCTGTGGAGCGCGA 9000
Db 8941 ccttctcccttgtcttctggctcccggaacttggctccctacactgtggagcgcg 9000
QY 9001 GCGACAGTGGCGGGGAAGACGTAGGCTCCGCCCGCGCTTGGGGCTTCCCGCGCGCG 9060
Db 9001 gcgacagtggcgcggaagacgtaggctccgcccggccttggggttcccccgcgcg 9060
QY 9061 CCGAGGCGCGCTCCCGCGGCGCTCTCCGGACTGGCGGTGGGGCATCCNCGGCGCG 9120
Db 9061 ccgagggcccgctcccgggcgctctcccgactgcgactggcggtgggcatccnngggcg 9120
QY 9121 GCCCGCGCGCGGCTTCAGCGCGCGCGCGCTTCAGCGCACGCGCGCGCGCGCGCG 9180
Db 9121 gcccgcccgcggttcagcgcccgcccgcggttcagagccacggcgcccgccccg 9180
RESULT 3
US-09-948-941-529
; Sequence 529, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000788
; CURRENT APPLICATION NUMBER: US/09/948, 941
; CURRENT FILING DATE: 2001-09-10
; PRIOR FILING DATE:
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 529
; LENGTH: 11190
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-529

Query Match 98.5%; Score 9040.4; DB 5; Length 11190;
Best Local Similarity 99.8%; Pred. No. 4.2e-169;
Matches 9165; Conservative 0; Mismatches 7; Indels 11; Gaps 11;

QY 1 CTGCTCTGAACCTCGGCTCAAGCAATCCTCTGCTTCAGCTTCCCAAGTGTCTTAA 60
Db 322 ctggctctgaacctcctggcctcaagcaatcctcctcgtcttcagcttcccaagtggttaa 381
QY 61 TTACAGGCATGAGCGCTGGCATGAACCTTGACACTATTGAGATATACTGGTCAGGTATT 120
Db 382 ttacagcatgagcctggcatgaaccttgacactatgagataatactggtcaggtatttg 441
QY 121 TGGAACTCCCTCAACTCTGTTTGGCAGATGTTTCTCATGATTAGGAGAGATTATAA 180
Db 442 tggaaatccctcaactctgtttggcagatgtttctcatgattagaggagattataa 501
QY 181 ATTTTGAAGAAATCCAGAGAGGTGAAGGTGAAGTGAAGGACAGAAATTAATCTGTTT 240
Db 502 attttgaggaaatccagagaggtgaagaggtgaagtagggcagaaatttaactgtttt 561
QY 241 ATTTTACTGCTATATACCGAGTGTCTGGAACCTTGGCCCATGTTAGTACCAAAAATCTGTT 300
Db 562 atttaactgctataaccagagtgctggaacttggcccatggttaagtaaccacaaatcgtt 621
QY 301 TTTTGTGAATGAATAAGCAATAAATGAGTCACCGTGGAAATTTAGTATTATTTTCAAGT 360
Db 622 tttttgaaatgaataagcaataaatgagtgacctggaaatttagtatattatttcaagt 681
QY 361 TTCAAAGCGTGTGATTACAGGCGCAGGACAGTGGCTTCACACCTGTAATCCACACTTT 420
Db 682 ttcaaagcgtgtgtgattacaggcgccagggcacagtggtcacaacctgtaatccccacattt 741

QY 421 TGGAGGCCGAGGTAGGAGGATCACTTGAAGTCAAGAGTTCGAGACCAGCCTTGACCAACAT 480
Db 742 tggaggccgaggtlaggagagatcacttgaggtcagaggttcgagaccagccttgaccaaat 801
QY 481 GGTGACACCCCTGTCTCTACTAAGTAAATACAAAATTAGCCAAAGTGTGGTGCCAGGCA 540
Db 802 ggtgacacccctgtctctactaagtaaaatcacaaaattagccaaagtgtggtgagcgca 861
QY 541 CCTGTAATCCCGGCTACTTGGGAAGCTGAGCGAGAAGAACATCACTTGAACCTGGGAGGCA 600
Db 862 cctgtaatcccgctacttgggaagctgagcgagaagaatacacttgaacctggagagcag 921
QY 601 AGGTGTCAGTGAAGCGGAGATCACCCACTGCTCCAGCCTCAGTGAAGTGAAGTGAAGTGA 660
Db 922 aggttgcagtgagcgagatcaccccaactgcaactccagcctgagtgacagagcgagactc 981
QY 661 TGCTCTCAAAACAATAAACAATAAATACTACTTGTGGCGGGTAAAGTGGTTCACGCGCTGT 720
Db 982 tgtctcaaaaacaataaacaataaactactcttggcgggtaagggtgggttcaagcctgt 1041
QY 721 AATTTTAGCACTTTGGGAGGCTGAGCGGCGAGATCACTTGAAGTTCAGGGGTTTCGAGACC 780
Db 1042 aattttagcaactttgggaggtcagcgcgagatcacttgaaggttaggggttcgagacc 1101
QY 781 AGTGTGCCCAACATGTTGAACCCCATCTCTACTTAAATACAAAAGTGTTCGCGTGT 840
Db 1102 agtctgscacaactgtgtaaccccatctctactaaatacaaaaagtgttcttctgggtgt 1161
QY 841 GGTGGCGGAGCGCTATTAATCCAGCTACTTGGGACTTTTTTTTAAAGCGGAATCTCACTC 900
Db 1162 ggtggcgagcgctataatcccaactcttgggacttctttaaagacgggaatctcactc 1221
QY 901 TGTGTCAGGCTGGAGTGGAGTGGAGTCTGGCTCACTGAAGCTCCGCGCTCCAG 960
Db 1222 tgttgcacaggtcggagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1281
QY 961 GTTCAAGGGGATCCCGCGCTCCAGCTCCCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGA 1020
Db 1282 gttcaa-gggattccccgcgctcagcctcccaagtagtgagggaatccctgtctctgcaa 1340
QY 1021 AAAAAAAAAAAAAAAAAAAAAA-ATATATATATATATATATATATATATATATATATAT 1079
Db 1341 aaaaaaaaaaaaaaaaaaaaaataataataataataataataataataataataataataata 1400
QY 1080 GTGTGTGTGTGTAT 1139
Db 1401 gttgt 1460
QY 1140 GGAGTGTGGCGCACGCTGTGATCACAGCTACTCGGAGGCTGAGGACGAGAAATCGCT 1199
Db 1461 ggaagtgtgagcgacgctgtgacacagctactcggagaggtgagcgagagagagagag 1520
QY 1200 TGAGCCCGTGAAGTCGAGGCTGAGTGAAGCCAGTGAAGCCACTGCAATTCACGCTGGG 1259
Db 1521 tgagcccgtaagtcgaggtgagtcgagtcgagtcgagtcgagtcgagtcgagtcgagtcg 1580
QY 1260 CGAAAGAGAAAGCGGTCTCAAAACAAACAAACAAAGCTACTTTCAGCACGTTAG 1319
Db 1581 cgaaagagaaagcgggtgtctcaaaacaaacaaacaaagctactctcttagcacgtgttag 1640
QY 1320 AGTATCTCGCGGGGGAAGTGGGAACAGTGTGTCACACAGTAGGATCATCTTTTATATG 1379
Db 1641 agtatctcgcgggcggaagtgggaacagagtgctgacacagagtaggacatcttatalg 1700
QY 1380 TTAACACACACTGATACCAGCTAAAGCGGTGAACACATTTACTCTCTGCGAGTGTATA 1439
Db 1701 ttaacacacactgataccagctaaagcggctgaacacatttactctctgagcagtgatta 1760
QY 1440 AAGTATCTGTTTTCATATTTTATTTTATTTTATTTTATTTTTCGATCAACCAACCTGAT 1499
Db 1761 aaagtaactgttttctcatatttttcttcttcttcttcttcttcttcttcttcttcttct 1820
QY 1500 CTTTTTCTCATACTTGGCCAGCGCTGACAGCAAAACCGGACAGAACTCGCGAC 1559

||||| 6200 gcagcgagtgagggtgggagtgagagtagagagactccctgggatcttctgtggcccc 6259
QY TTCTGGGTGTCCTCCCTGGTGGGGCATTTCTGCGCACAGGCGAGCTGGGGCTCCCTCCCTGA 5998
Db ttctggggtgtgccccgtgggtgggagcatgtgtgcagcagggcagctgggggtcctccccga 6319
QY GGATCCTCTGGCTCACCTCCATCCATCCAGGCGATTGCTCAGCCAAAGACCTACTATCGGGATG 6058
Db ggatcctctgctcaactccatccatccagggcatgtgctccagcaagacactactatcgggatg 6379
QY AACACATCTACCCCTPACATGTACCTGGCTGGCTACCACTGCGCAACCGCAATGTGCGGG 6118
Db aacacatcaccctacatgtacctggctggctaccactgtcgcaaccgcaatgtggcgg 6439
QY AAGCCCTCGAGCCTGGGGCGACAGCGCCACTGTCTACCCAGGATGAGGATCCCCCTACT 6178
Db aagccctcaggcctggggagacggcgaactgtcatccagagtgagga tccccctact 6499
QY AGGGCTCGAGCCTGTCTTTCTTCCCTCCATCATAGTTTCCAAACACACCTCGTCCAGGAC 6238
Db agggcctcagcctgtctcttctccctccatcagtttccaaacacccctcgtccaggac 6559
QY TGAGGCCCTGGCTCCACGCCCCATCCCTTTCCATPCCATGCTCCCTAGGCGAAGGCCACC 6298
Db tgaggcctggctcccaagccccatccctttccatccagtcctcagtcagcagcaagccacc 6619
QY ATTACCCAGGAGGTAGGACCCCTGATTAAGGTGTACATCTTCCCTCCCTCCCTCTCC 6358
Db attaccaggaggtaggacccctgattaaggtgtcaactcttccctccctccctctcc 6679
QY TCCTAATTTTTTTTTTCTCAGAACAGTCTCAAACTCCCAATGTTTAAACACCATCATCCA 6418
Db tcttaatttttttttctcagaacagtcataaattccaaatgtttaaacacacatccca 6739
QY GCAGTGGACTTCCACCTCGGCCCCATGCCCCCCTCTCATCTTGTCTTTCCTCTG 6478
Db gcagtgggacttccacccctggccccatggccccctctctctctctctctctctctctg 6799
QY GCCTGACCCAGACATANTTTGACGTAGGAGCCACCTACTCCCTCCCTCCCTCTGG 6538
Db gctcaccacagacagca tcaatttgcagtggagagccccacactactcccccgccccctgg 6859
QY GSCTCCATCCCGCCAGGTCTCGGGCTACCCCGCATGCTGAGACCCCTTCAGACCT 6598
Db ggtccatcccccgccaggttccctgggtactaccctccagatgtgagaccccttcagacct 6919
QY ACAGAGCCCTACTCTCTACAGCTACAATCTGCGGGGAGAGAGGAGATCTACAA 6658
Db acagagacccactgtctcacagctacaaactactgcggggagagagagagatacaaa 6979
QY GGAGTTCTTTGAAGTAGCAATGATGTATCCCAACCTGCTGAAGGAGGAGCCAGCTT 6718
Db ggaagtctcttgaagtagccaa tga tgcatacccccaactgtctgaaggaggcagcagctt 7039
QY GCTGAGGCGGGCGAGGAGCGCGGGGAGCAAGCCAGGTGAAGGCTGGAGCTCCAG 6778
Db gctgagcgggcgagagcgcgcggggagcaagccaggtgaaaggctggagctccag 7099
QY CTTGTGTCAGCCTCCACCTGAGCAGGGCTCCCTTCCACAGGGCCATGGGGCTGTCATG 6838
Db cctgtgtccagcctcccaactggacagggtcccttccaca -ggccatgggggtgcag 7158
QY TACGGGATTAGGAGTAGGAGGAGGAGGTGGCCCTGAGCAGACGCTATGTTCCCTTTTG 6898
Db tacgggatttagggatggcaggaggaggtggccccctgagcagacagctatgttccccctttg 7218
QY CTATAACTGAGTCTGGGCCACCTTGGACGGGACTGAAGGTATTTAGAGGTTTCTAC 6958
Db ctataactgaggtcctggggccccagcttggacgggactgaaggtatttttagaggtttctac 7278
QY CTTGTGCTTACGTTTCATGGCCAGACTTCCCTCCCTCAGCTGAGGGGTAGGAT 7018
|||||

Db cctgtgcttccagtttcatggccagactccctccctcagctgaggggtgagggtagggat 7338
QY GGTACGTCTCGCTATGATTGGCTTTATAAAGAAAGAGAGGTCTTAAGAAATGTTCCCAA 7078
Db ggtacgtctcgtctatggatttggctttataaaagaaagaggttctaaagaaatgtlcccaa 7398
QY CCTATGCTTACTTTTCTGGAGCCAGGGGTCTTTGCCCTAGGTGGGGGCTTGGCTCTGTC 7138
Db cctatgcttactttcttctggagccaggggtcttctgctcagtgaggggtcctggcctgctgc 7458
QY CCTCTGTAAGGGGTAGTAAGACACTGATCTGTGCCCTCCCTTCCCTCCCTCGTCAGGGC 7198
Db cctctgctaagggttgagtaagactgatctgtgccctcccttccctcctcgtccagggc 7518
QY ACCCAGAGCCAAAGTTCCGCCCTCCAGGACCTGAGTCTTGGCCACTCTGCTCGATTTC 7258
Db acccagagccaaaggttccgcctccctccagaccctgagtgcttgcgccacactcgtgcgaltc 7578
QY TACGACGSCATCTCAAAATGGGAGGAGGCGAGTCCCAGCCCTGTCTGCACGTGGGCTGG 7318
Db tacgacggcatctgcaaatggagagggcagtcctccacgcctgtgctgcagtgggctgg 7638
QY GCCACCTTTCTTGTGCAGTCCCTAGGCCGTTTTGAGGGACAGGTGAGGACAGCTGCACA 7378
Db gccacctttcttctgcatcctcagtcctgtggagacaggttgagggacagctgcaca 7698
QY GAGGTCTGGGCACTACAGTGTGTGACAGCACGCGGTGTGTGAGACTTTTCTGGCCCA 7438
Db gaggtctggggcaactacaggttggtagacagcagccagggctgtcaggaactttcttggccca 7758
QY GGGCAGCAGTCTGCCATPCCTCTTGGTGGGCTGGGACTGAGACCCCTCGGTGGGATG 7498
Db gggcgagcatctgccccctctcgttgcagtgaggactgagacccctggg lgggagtg 7818
QY GGATGGCCAGAGCGGTCTTGGAGTTCCAGCCACTGCCCGGCAACCTTGTCTCACCTT 7558
Db gga tggccagagcaggggtccctggagttccagccactggccggcaacctgtctcacctt 7878
QY GCTCTCCCTACTGGCCAGGTGCGGCAAGGTGCGCATAGTGAGCCGAGAGCGCCAGGC 7618
Db gctctccccactggccccaggtgcggcagaaaggtgcgcatagtgagccagagagcgagcg 7938
QY GCGCAGGCCGAGGAGCGCTGGGGCGAGGAAGCCCGGAAAGCCGGCGGGCGGGCCACG 7678
Db ggcgagggcgagagcgctggggtggaggaagcccggaagcgcggtggggggccccag 7998
QY CGGGAGTCCAAAGCAGAGGAGCGCCCGCCCGCAAGAGCCAGCACTGGACAGGGCTT 7738
Db gggggag tcccaagccagagggagcccccgcccccaagagccagcactggacaagggcct 8058
QY GGGCACCGCCGAGGCTGAGTGTGAGGACCCCGGGAAGCTCTCTGGGACTGTCTCGCTGG 7798
Db gggcagcgccaggggtgagtgtagagaccccccggaagcctcctctgggactgtcgctgg 8118
QY CACAGCCCGAGGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCAC 7858
Db cacagcccgagggcctgaaggtggcagcacggtcaggtggcagcagcccgagcagctacc 8178
QY ACCCGCGAGGCTCAGTGTCTACTTTCCAGAGTGAAGAGATGAAGGCAATGAAGAGCT 7918
Db acccgaggagggtccagtgctcaacttccagagtgagaga tgaaggagac lgaaggagct 8238
QY GCTGTGGCCCAACAGATCACTCGAGGCCATCAAGCTGAACCTGCAACCTCAGGGCACAGTCGA 7978
Db gctggtggccaccaaga tcaactcgagccatcaagctgcaactgcaagcagctgcag lcgca 8298
QY AGTCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTCTCTCA 8038
Db agtgagatgaagaagcagaaagtgtccacccctagtactacactctgtcttctctcaa 8358
QY GCGGAGCGCAAAAGCCCTGTGAACCTACTGGGACTTCCGACCGCTTGTGGGACCCAGGC 8098
Db gggggcagcgcaaaagccctcctgaactcctgggagcttgcggaccgctgtggggagcccaaggc 8418

Qy	9178	CCG 9180
Db	9494	cag 9496
RESULT	4	:
US-09-948-941-649		
; Sequence 649, Application US/09948941		
; GENERAL INFORMATION:		
; APPLICANT: VENTER, J. Craig et al.		
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED		
; WITH CANCER, METHODS OF DETECTION AND USES THEREOF		
; FILE REFERENCE: CL000788		
; CURRENT APPLICATION NUMBER: US/09/948,941		
; CURRENT FILING DATE: 2001-09-10		
; PRIOR APPLICATION NUMBER: 60/231,328		
; PRIOR FILING DATE:		
;		
; NUMBER OF SEQ ID NOS: 12618		
; SOFTWARE: fastseq for Windows Version 4.0		
; SEQ ID NO 649		
; LENGTH: 11205		
; TYPE: DNA		
; ORGANISM: Human		
US-09-948-941-649		
Query Match 98.5%; Score 9040.4; DB 5; Length 11205;		
Best Local Similarity 99.8%; Pred. No. 4.2e-169;		
Matches 9165; Conservative 0; Mismatches 7; Indels 11; Gaps		
Qy	1	CTGGTCTTTGAACCTCCTGGCCCTCAAGCAATCCTCTGTCTCAGCTTCCCAGAAGTGTTGTAA 60
Db	322	ctggcttctgaactcctggtcccaagcaatcctcctgcttcagcttcccagaagtgtgttaa 381
Qy	61	TTCAGGCATCAGCGCTGCGCATGAACCTTGACACTATTGAGATATCTAGTCTCAGGTATTTG 120
Db	382	ttaacaggcatgagcctggcgatgaaacttgaaactattgagatactagtgcaggtattttg 441
Qy	121	TGGAATGCCCTCAACTCTGTTTTGCCAGATGTTTTCTCATGATTAGAGAGAGATTATAA 180
Db	442	tggaaatgccctcaactctgttttgccagatgttttctcatgatattagagagagattataa 501
Qy	181	ATTTTGAGGAAAAATCCAGAGAGGTGAGAGGTGAAGTGAGGCGACAGAAATTTTAATCTGTTTT 240
Db	502	attttgagaaaaatccagagaggtgaaagaggtgaaagtagggcagaaaattlaacttgtttt 561
Qy	241	ATTTTACTCTATATACCAGAGTCTCTGGAACCTTGCCCATCGTAGTACCAAAAATCTGTT 300
Db	562	atttactgtcata lacogag tgtctgaaacttgccccattggttaagaccagaatactgtt 621
Qy	301	TTTTTTTGAATCAATAAGCAAATAATCAGTCACCGTGGAAATTTAGTATTTATTTCAAAGT 360
Db	622	ttttttgaaatgaataagcaataaatgagtgacctggaaaaattaglattattttccaagt 681
Qy	361	TTCAAGCGTTGTTGTATACAGGCCAGGCACAGTGGCTCACACCTGTGAATCCCAGCACTTT 420
Db	682	ttcaaacgcttgtigtatatcaggccaggcacagtggtccacacctgtaatccccagcattt 741
Qy	421	TGGAGGCCGAGGTAGGAGGATCATTGAGGTCAGGAGTTTCGAGNACCAACCTGCACCAACAT 480
Db	742	tggaggccgaggtgaggagatcacttgaggtcaggagttctgagaccagcctgaccacaact 801
Qy	481	GGTGACACCCCTGTCCTTACTTAAGTAAATAACAAAATTAGCCAAAGTGTCGTGGCAGGCA 540
Db	802	ggtgacacccccctgtctctactaagtaaaaaatacaaaaattagccaag tgtggtggcaggca 861
Qy	541	CCTGTTAATCCCGGCTACTTGGGAAGCTGAGCGAGAAAGTAATCATTGAACCTTGGAGGSCAG 600
Db	862	cctgtaatcccggtacttgggaagctgaggcagaagaataacttgaaccttggaggcag 921
Qy	601	AGGTTGCAGTGAGCCGAGATCACCCCTACTGCATCCAGCCTGAGTGACAGAGCGAGACTC 660
Db	922	aggttgcagtgagccgaatcacccctactcacttccagcctgaatgacagagcgaagctc 981

QY 661 TGCTCAAAACAAATAAATAACTACTCTTTGGCGGGTAAGGTGGTTTCAACGCTGT 720
|||||
Db 982 tgtctcaaaacaaaataaactactcttggcgggtaaggtggttccagcctgt 1041
QY 721 AATTTTAGACTTTGGGAGGCTGAGCGGGCAGATCACTTAGGTTAGGGTTCCAGACC 780
|||||
Db 1042 aatttttagcaatttggaggtgagcgagatcaacttgaggttaggggttagagacc 1101
QY 781 AGCTCTGGCCACATCGTGAACCCCATCTACTTAAATACAAAAAGTTTCTGGGTGT 840
|||||
Db 1102 agtcgggcacaactggtgaaccccatctactcaaatcaaaaagttttctgggtgt 1161
QY 841 GGTGCGGACGCTATAATCCAGCTACTTGGGACTTTTAAAGAGGAATCTCACTC 900
|||||
Db 1162 ggtggcgagcctataatccagctacttgggaacttttttaagacggaatctcactc 1221
QY 901 TGTTCGCCAGGTGAGTGCAGTGGCAGATTCCTGGCTACCTGAGGCTCCGCCCTCCAG 960
|||||
Db 1222 tgttggcaggtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1281
QY 961 GTTCAAGGGATTCCCGGCTCAGCTCCCAAGTACCTGGGAATCCCTGCTCTGGAA 1020
|||||
Db 1282 gttcaa-99ga tccccgcgcctcagctcccaagtagctgggaatccctgtctctgcaa 1340
QY 1021 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1079
|||||
Db 1341 aaaaaaataaataaataaataaataaataaataaataaataaataaataaataaataaataa 1400
QY 1080 GTGTGTGTGTAT 1139
|||||
Db 1401 gtgtgtgtgtatataatattatgtatgtatgtatgtatgtatgtatgtatgtatgtat 1460
QY 1140 GGAGTGTGGCGCAGCTGTGATCACAGTACTCGGAGGCTGAGGACAGAGATCGCT 1199
|||||
Db 1461 ggaagtggcgagcgcctgtgatacagctactcggaggtcagagagagagagagagagag 1520
QY 1200 TGAGCCGCTGAAGTGCAGGCTGACGTAGCCAGATGAGCCACTGCATTTCCAGCTGGG 1259
|||||
Db 1521 tgagccgtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgagtgag 1580
QY 1260 CGAAGAGAAAGACGCTGTCAAAACAAACAAACAAACAAAGTACTTTCAGACGTGTAG 1319
|||||
Db 1581 cgaagagaaagacgctgtcctcaaaacaaacaaacaaagctactcttagcagcgtgttag 1640
QY 1320 AGTATCTCGCGGCGGAAGTGGGAAACGAGTGTGCACACAGAGTAGGACATCTTTATATG 1379
|||||
Db 1641 agtatctcgggcggaagtgggaaacagtgctgcacacagagtaggcatttctatatg 1700
QY 1380 TTAACAGACACTGATACCCAGCTAAGCGGCTGAACACATTTACTTCTTGGCAGTGTATA 1439
|||||
Db 1701 ttaacagacactgatccagctaaagcggtgaacacatttactctctgagtggttta 1760
QY 1440 AAAGTATCTGTTTCTCATATTGTTTATTTTAAATTTTCTGGATCAAGCAACTGAT 1499
|||||
Db 1761 aaagtctctgtcttctcata tgttttatttatttatttcttcttcttcttcttcttcttcttct 1820
QY 1500 CTTTTTCTCATAACTGCCAGCCAGCCGTGACAGCAAAACCGCAGAAAGCTCGCGCAC 1559
|||||
Db 1821 ctttttctcataacttgccgacgacccgtgacagcaaaacccgagagagctcggcag 1880
QY 1560 CTCGCCACCCGAGTCTGAGGTAGTGGCCCGGACTACATTTTCAGAGGACTTCGCGGG 1619
|||||
Db 1881 ctccaccccgagtgaggttagtgcccccgaactatcttccagagggcaactgaggg 1940
QY 1620 CACGCTTCTGCTGCTGGCTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1679
|||||
Db 1941 cagcgtctctgctgctggctcagggagggagggagggagggagggagggagggagggaggg 2000
QY 1680 GGTGTCCGAGCCGCGGACCTAGAGATCCAGAGCCACAGCGCGCGCGCGCGCGCGCGCG 1739
|||||
Db 2001 ggtgcccggagcgcgacctagagatcccagaagccacagcgagcgccgcccgc 2060

QY 1740 ACTATTTCCAGCTCTGCGGGCAGGGTGGGCCCCAGACTCCACATTCGCCGCGGTAGTG 1799
|||||
Db 2061 actatttccagctctcgggcaggggtggccagactccacttcccggggtagtg 2120
QY 1800 CGACCTTAGGGCGGACTTATGTCCAGAGGCTTCGGGGCGGCTGCGCGCGGTGCC 1859
|||||
Db 2121 cgacctagggcgggacttcatgtcccagcaggtcccgggcggtgctgcgcgagtgcc 2180
QY 1860 TAGTGTGGGATGTAGCGCGGAGGTGGCGAGGGGAGCCGAGCCAGGACTCTCTTCTGGG 1919
|||||
Db 2181 tagtgtggagtgaagcgcgagggtggcgaggggagccagggagactctcttggg 2240
QY 1920 GTTTGGGGCTTGACTGGGTGCGCTTTCTGCACAGACTTTACAGCCCGGGGACAG 1979
|||||
Db 2241 gtttggggcttgacctgggtgctcttctggacagactttacag-ccccgggggacag 2299
QY 1980 TCGTAGAGAGGGGGGGGGGCCCATTTGGGGCTCTCATTTGGGGTCTTTGGGGCGCACCC 2039
|||||
Db 2300 tcttagagagggcgggcgccatttgggtctctctatgggtgcttggggcgaccc 2359
QY 2040 CATCGGTACCGGGCTCCCGAATTTGTGGGGACAAAGGCTCTGCAGTCTCGCTGA 2099
|||||
Db 2360 catcgggcaccggggcgctcccggaatctgtgggggacaaagagctctgcagctcggctga 2419
QY 2100 GGGTCTCACCGACAAAAGAGGGGAGCGGTGAGCAGAGGCTGAAGAGGGTGGCGAAGC 2159
|||||
Db 2420 ggggtctcaccgacaaaagaggggaagccggtgagcagaggtggaagaggtgggggaagc 2479
QY 2160 AGGGAGCTGTGCTGTGTCGGGGCGGGTGGAACTTTAGCGGACCTTGGGAGGAGGCTCC 2219
|||||
Db 2480 agggagagctgtgctgtgctggggcggtgtgaaccttagcgagacctgggagagagctcc 2539
QY 2220 CCGCGCAAACTGCCCGACCTCCCTCCCGCGCTTCTGACGGCGCGCCCGCCACCGC 2279
|||||
Db 2540 ccggcgaaactgctccgacacctccctcccccggtctgcttgcagggcgccgacacgcg 2599
QY 2280 CCGCGCCCATGGGGCTGAAGCGCCCGCCAGAGAGCTGTTCCTCGCTGCCCTCCATCGAGC 2339
|||||
Db 2600 ccgcgcccatggggtgaagcgcccgccagagagcgtgttccccgctgcgctccatcgagc 2659
QY 2340 ACGTGTGCGGCTGTTTCTGCCGAGCTGGCCGAGAGGAGCGGACCTGTGTCTCTTT 2399
|||||
Db 2660 acgtggtgctgctgttctgctccgagctggcgagagagagagagagagagagagagagag 2719
QY 2400 CTTTGGTCTGCTGGCTTCGTGGAGCTTTTCTGGCTGTCAACCGCTCATCTCCATCAACAG 2459
|||||
Db 2720 ccttgggtgctgggttctgagcaatttcttggctgtcaaccgctcatctccatccacag 2779
QY 2460 TTCCGAGCTCACCTTCAGCCCGCGCCCGCCCGAGCGGCTGGGGCTTCACCTACT 2519
|||||
Db 2780 ttccgagctcacctccagcccgccccgaccccgagcctggtgggctcaactact 2839
QY 2520 TTCCGCTGGCGACCTGTCTATCATCGCGCGCTCTATGCGCGCTTACCGCCCGAGATCC 2579
|||||
Db 2840 ttcccggtggcgagcgttctatca tgcgcgctctctatgcgcgcttcccccagacc 2899
QY 2580 GAGCGCGCTGACCTGTCTCTATCTTCGAGAGAGGGGGTGTCTTCCAGCGGTGAGCTGG 2639
|||||
Db 2900 gaggcgctgacactgct 2959
QY 2640 TGAAGAAGGCTTCGATGTATGGAACAGCCTCAGCGGCTCTACTTTCAAGGATCGGG 2699
|||||
Db 2960 tgaagaaggtctcagtgctata tggaaacgctcagcgctctctacttcaagggtcggg 3019
QY 2700 CCCACATCCAGTCCCTCTTACGTTTCATCAGAGTTGAGCCCGCTAGCTGGGAATCTTA 2759
|||||
Db 3020 cccacatccagctccct 3079
QY 2760 TCCATGACCCACTTCTTCAAAACCCCTCCATGTTTACAGAACCCCTTTTAAAGAACTGTAG 2819
|||||
Db 3080 tccatgacccacttcttcaaaacccctctctgttttcaagaaccccttcaagaactgttag 3139
QY 2820 CTTGTGAGGTTCCGCGAGGTGTATTTTCTCTCTTGTGAGTGGGAAACTGAAGCCACAG 2879

|||||
Db 3140 ccttggaggttcgaggtgttatcttctcttcttgcagttgggaactgaagccagag 3199
QY 2880 AGGGGAAATGATATGCAAAAGTCACACACGGCATGCGAGGGCTGGAAGTGAAGCTGATC 2939
Db 3200 aggggaaaagtatgcacaaagtcacacacgacatgacgaggtcggaagtgaaagcctgac 3259
QY 2940 ACTGGCTCCAAATCATCAACCTCACCTCTGCCCCCTCAGCACCCCCACCCCTTGGCACTG 2999
Db 3260 acttggctccaaatcatcaacctcaacctgtgccccctcagcaccaccaccttgccactg 3319
QY 3000 AACAGCTACAGGAGTTCTAAGCATGAGACACAGAGGGCGGCACAGATTTAGGGGGCAAG 3059
Db 3320 aacagtcacagaggttctaagcatgagacacagagggcgagcatgttagggggcaag 3379
QY 3060 AAGATGAATTTGGCTGCAATTTGAGGCAGTTAAACGAATAAATGGCTATCAAGATTTTT 3119
Db 3380 aagatgaatagggtcgatcttgaggcagctaaacaaaaataaggctatagaagattttt 3439
QY 3120 TTTTTTTTTTTTTTGGAGACAGGCTCTCACTCTGTCCCCAGGCTGGAGTGCAGTGGTGT 3179
Db 3440 ttttttttttttttggacaggtctcaactctgtccccagagctggagtgcaagtgtgt 3499
QY 3180 GATCATGGCTCACTGACGCTCAGTCTCCTGGGCTCAGAGATCTCCCAACCTCAGCCTC 3239
Db 3500 gatcatggctcaactgagcctcagtcctcctgggtcagagatcctccaaacctcagcctc 3559
QY 3240 CTCAGTAGCTGACAGTACAGGCATGACCGTGTGCTGGTAAATTTTGTATTTT 3299
Db 3560 ctgagtagctgagagtagagggcatgacaggtgtgtgtgtgtgtgtgtgtgtgtgtgtgt 3619
QY 3300 GTAGAGTGGTGTCTCACTATGTGGCCGACACTGGTCTTGAACCTTGGGCTCAAGTGAT 3359
Db 3620 gttagataggtgtctcaactatgtggccagactgtgtgtgaactttgggtcaagtgtat 3679
QY 3360 CTGCCGCTCAGTCTCCCAATGTCTGGGATACAGGTGTGAGCCACCGCAACTGTGTGGC 3419
Db 3680 ctgcccgcctcagctcccaaatgtgggtattacaggtgtgagccaccgcaactgtgtggtg 3479
QY 3420 CTATGAAATTTTTTTTTTTTTCAGAGGCGTCTCACTCTGCGCCAGGCTGGAGTGC 3479
Db 3740 ctatgaaaatttttttttttttttcagacgggtctcaactctgtcgccaggtggaggtgc 3799
QY 3480 AGTGGCTCAATCTCGGCTCACTGCAAGCTGTGCTCTCTGTCTATGCCAATCTCCTGCC 3539
Db 3800 agtgggtgcaatctcggtcaactgcaagctgtcctcctgtttcaatgcattctcctgctg 3859
QY 3540 TCTGCTCAGCTCCTGAGTACTGGGACTACAGGAGCTGCGCACCATGCCCTGGCTAAT 3599
Db 3860 tctgtcctcagctcctgtagtctggactggactacagagagctgccaacctgctggctaat 3919
QY 3600 TTTTTTTTTGATTTTAGTAGACAGAGGTTTCCACCATGTTACCGAGGATGGTCTGATC 3659
Db 3920 ttttttttgatttttagtagaacagaggtttcaccatgttagccaggttaggtctcgatc 3979
QY 3660 TCTGACCTCGTATCGCCCGCTTGGCTCTCCCAAGTCTGGGATACAGGCGTGAGC 3719
Db 3980 tcttgacctgtgatccgcccgttggtcccaaatgtgtgtgtgtgtgtgtgtgtgtgtgtgt 4039
QY 3720 CACCGACCTGGTCAAAATGTTGAGACAGAGAGGGGCTTGACCTCAAAAGGCTTAAG 3779
Db 4040 caccgacacctggccaataatgtttgagacagagaggggttgacctcaaaaggtctaaag 4099
QY 3780 ASTCAGGGCTTGAAGAGCTTTCACCAAGCCGCTTGACTGGCAATCCCATCTCGGTG 3839
Db 4100 agtcagggcttgcaagagagctttgaccgaagcccggttgactggcaatcccaacctgggtg 4159
QY 3840 TGCCATATTGAGAGGAATCAGAGGCTGTCTCTCAGCTTAGCAGGAAAAGAGTGCAGAGA 3899
Db 4160 tgccatatgagaaggaatcagaggtgtcttctcagcttagcagggaaaaagagtgacagaga 4219
QY 3900 TAAATGAGGGTTATTTGTTGGTGGGTGTATAGCCAGAGAGTGTGGCCAGCGTCTGTTT 3959
|||||

Db 4220 taaatgaggggtattattgttggtgtgtatagccagagagtggttggccagcgtctctgtt 4279
QY 3960 TTGCCATTCCTGTTTTAACTAGTAGTCAGTAAATGAAATGCCCTAAATCCATCAATAAT 4019
Db 4280 ttgccattctctgttttaacctagtaagtcagtaaaatggaaatccctaaatccatagaat 4339
QY 4020 ATATAATAGAGTTGCGAGAGAAAGACGAGTAGGGCCAAAGGCTGGGTACGTACAGGATA 4079
Db 4340 atataatagagttgcagagaaagacagaggttagggccaaaggtcgggtcagctacagata 4399
QY 4080 TCCAGAAAGGTATCTTGTGGACATAGAGGGTGTAAACACAGGAGAGAGTCTTTCAACACG 4139
Db 4400 tcagaaaaggtactgtgttgacatagaggtgttaaacagggagagaggtctttgaaacag 4459
QY 4140 TGGAGGGAAGGATGAGGATAGTGGCAGGAGATCTGAGGTGGGTGCACAGGCTTG 4199
Db 4460 tgggaaggaaggtatggagaggaatagtgggcagaggaatctgaggttggttcaacagcgtg 4519
QY 4200 GAAAGGAGTGGAGGAGTGTGGCCCATCACTACCTTGG-CCCTTTTCCCCTATGTTAAAG 4258
Db 4520 gaaagggagtgggagggagtgtggcccatcactacctggccccctttccccatgttaaag 4579
QY 4259 CACAGAGGACCTCTTTCATTTACTTCCCTTCCACAGGACCAAAATGGACAGCTCCGG 4318
Db 4580 cacagaggacctctttacttacctcccccttcacaggcaccaaattggacagctcgg 4639
QY 4319 TGTGGCTTTGCTGTGGTGGGCTGCCAGGCGCTGGGTCTCGGGATGTCCACCTCGC 4378
Db 4640 tggggcctttgtgtgtgtgtggcctggcagggccctgggtctccggagtgtccacctcgc 4699
QY 4379 CCTGCTGAGGATCATGCCCTGGGTAGTGTGGGCCCAATGGGAGCAGACAGCTGAGGT 4438
Db 4700 cctgtctgaggtatcgtctgggttagtgttttggcccaatggggagcagacagctgaggt 4759
QY 4439 CACTGGCAGCGCAAGGGCAACGAGGACCCAGGGGCCACACAGTCAATGCCGTGTGGC 4498
Db 4760 caoctggcagcggaaggcaacgagacggcaggggccaacagtcgaatccgggtgtggc 4819
QY 4499 TGAGCGGTTATTGCTCCCTCCCGCAGCTGTGCCCTTCTACTGTAGTACGCCAAGCC 4558
Db 4820 tggagcgggtattgttccctccccccagcctgttcccccttcatactgtatgagcccaagcc 4879
QY 4559 ACCCAAGGACTCCATTTTCTTGGGCCACACCCCTTTCTTCCCATCACACCCACATAGG 4618
Db 4880 acccaagggactccatttcttggccacacccctttcttcccatcaccaccaLag 4939
QY 4619 AAGGGAAGACAGAGACCCCTTTCTGCTGTGCTATTCCTGAAGCAGGACAGGCTGG 4678
Db 4940 aagggaagacagaagagcccttttctgggtgtgtcatctccctgaagcaggggtgg 4999
QY 4679 GCCATCATGAGACATAATGATCTCATCCCCCTTAAGAGCTGGCTGACCTGAAAGATC 4738
Db 5000 gccatcatgagacataatgatctcatcccccttaagagctggctgLaactgaaagatc 5059
QY 4739 ATACATCGCTGTGACCGCAAGATGAGGTGGCTTTCATGTGTGTGTCCTCAACCCCTTC 4798
Db 5060 atacatgctgtgacccgcaagatggaggtgggttcatggtgtgtgtgtgtgtgtgtgtgtgt 5119
QY 4799 CATTGACCTGCACACCACGCTGCTGGAGCTTCTGCAGCTGCAGCAGGTGAGGGCTGAGCC 4858
Db 5120 cattgaactgcaacccgactcgtggagcttctgcaactgcaagcaggtgaggggtgagcc 5179
QY 4859 AATGGGCGAGACTGGGCTAGGCCAGACTTGACTTGTGTGGGACCTGGCGACGGGCAC 4918
Db 5180 aaLggggcaggaactgggtcagactgactgtgtgtgggacctgggagacctgggagggcac 5239
QY 4919 TTTCCCTTCTGAGCTTCAGCTTCCCTCCTCGGAAAATGGGTAGTAATTCCTGGGCTG 4978
Db 5240 ttccccctctgacttgcagcttccccctcctcctggaaaaatgggttagLaaLttcctggcctg 5299
QY 4979 GCCTTTCCCGAGGGCTCTTGGGAGAGTAGAATTGAGATGTGAAATTCCTTTGACTTCCATTA 5038
Db 5300 gcctttccccgggctctcttgggagagtagaattgagatgtgaaattgtcttgaactccatta 5359
|||||

||||| 7519 accagagccaaaggttcgcccctccaggccctgagtgttcgcccacacctgctgcgattc 7578
QY TACGACGGCATCTGCAAAATGGAGGAGGCGAGTCCACACGCTGTGCTGCACGTGGGCTGG 7318
Db taccagcggcatctgcgaaatgggagggggcagtcgccacgcccctgctgcgcagctgggctgg 7638
QY GCCACCTTTCTTGTGACGTCCTTAGCCCTTTTGTAGGGACAGGTGAGGACAGCTCGACA 7378
Db gccacctttctgtgcagtcgccctagccgttttgaggagcaggtgagggacagctgcaca 7698
QY GAGGCTGGGCACTACAGTGTGTGACAGACGACCGGCTTGTACAGCTTTTCTTGCCCA 7438
Db gaggtctggcactacaggtgtgacagcagccagcggtgttcagacttttctggccca 7758
QY GGGGACGATCTGCCATCCCTTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7498
Db gggggagcatctgcccaccccccttcggtgcccagtgagacccccctgggtgggagtg 7818
QY GGATGCCACAGCAGGCTCTCTGAGTTTCCAGGCACCTTGGCCGCAACCTTGTCTCACCTT 7558
Db ggaatggcagagcaggttccctgagttccagccactggcggcaaccttgcctcaactt 7878
QY GCTCTCCCACTGGCCCAAGGTGGCGGAGAGGTGCGCATAGTGTAGCCGAGAGCCGAGGC 7618
Db gctctcccacttggcccaggttgcggcagaggttgcgcatagtgagccgagagggccgagtc 7938
QY GCGCAGGCGGAGGAGGCTGTGGGCGAGGAGCCCGGGAGGCGCGGCGGGGCCACG 7678
Db gggccagggccgagggagcgctggggcgaggaagcccgggagggcgggcgccacg 7998
QY GGGGAGTCCAAAGCAGAGGAGGCGCGCGCCGCGCCGCGCCGCGCCGCGCCGCGCT 7738
Db ggggagttccaaagcagagagggcccgcccgcccaagaagcagcaactggacaaagggcct 8058
QY GGGCACCGGCGCAGGTGTGAGTGTGAGGACCCCGCCGGAAGCTCTGTGGACTGTCTGCTGG 7798
Db gggcacccggcaggggtgcaggttgcaggaaccccccggaagcctctctggactgtcgctgg 8118
QY CACAGCCCGAGGCGCTGAAGTGGCAGCAGCTGAGTGGCAGCAGCTGAGTGGCAGCAGCT 7858
Db cacagcccgagggccttgaggttggcagcagcgtcaggttgcagcagcccgagcatcac 8178
QY ACGCCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGAGTGAAGGCGATGAAGGAGCT 7918
Db accgcccaggggtccagtgctcaacttccagagtgaagagatgaagggcagtgagagct 8238
QY GCTGTGTGCCACCAAGATCACTCGAGCGCCATCAAGCTGCAACTCACGGCAGAGTCCGA 7978
Db gctgtgtggccaccaagatcaactcgagcgccatcaagcttgaactcagggcacagtcgca 8298
QY AGTGACAGTGAAGACGACAAAGTGTCCACCTCTAGTGACTACACTCTCTTCTCTCAA 8038
Db agtgcagatgaagaaagcagaagtgctccacccccctagtgactcaactctgtcttctccta 8358
QY GCGCAGCGCAAGGCTCTGAACTACTGGGACTTTCTGGACCGCTTGTGGGACCCAGGC 8098
Db gggcagcgcaaaagcctctgaactactggggaacttcgagcagcttgtgggacccaggc 8418
QY TCGG-CCTTAGTCCCACTCTGAGCCCATGTTCTGCCCCCGACGCCCAAGGGGACAGGC 8157
Db tcgccccttagtcccccaacttgagcccatgttctgccccagcccaaggggacaggc 8478
QY CTCACCTCTACCCAAACCTAGTCTCCCGTCCCGAGTACGCTCTATCAACCCACGA 8217
Db ctacactctaccaaaccttaggttccccggttccccgagtcagtcgtatacaaacccacga 8538
QY TTTTCTCCAGCTCAGAACCCAGGCTCTGCCCCAGTCTGTAGAAATAGGCTCTCTCTCC 8277
Db ttttctccagctcagaacccagggctctgccccagtcgtagaataaggtctctctccc 8598
QY CAGAATCCAGCGCGGCAATGGAACCTCACGCTGGGTCTTAATACCAAGTCTTTAAAG 8337
|||||

Db 8599 cagaatcccaagcggccaatggaaaacctcacgctgggtcctaataaccagttcttaag 8658
QY CCCAGCCCCCTAGAAACCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAAGCT 8397
Db cccagccccctagaaaccccaagctctctcctcggaaccgctcacctagagccagaccacgt 8718
QY TACTCAGGCTCCTCCAGCTTGTAGAGCTGAGGTTTACCCCTTAACCCAAAGGAGCAC 8457
Db tactcagggctcctccagcttgttagagctgaggtttcaccccttaacccaa-ggagcac 8777
QY AGTCCCACTCCAGCCCGGGAGCCTAGGACCACTAGCCCTAGGAGTATATTTCCCGC 8517
Db aggtcccaactccagccc-gggagcctaggaaccaactcagccccctagagtatattccgc 8836
QY ACTTCAGAATTCATATCTTGGCAATCCCAAGCTCCCTGCCCCCAAAATAACTTTCAGTCT 8577
Db acttcagaattccatactctggaatcccaagctccccctgccccaaataacttcagtcctgc 8896
QY TTCAGAATTTGGAATTCCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8637
Db -tcagaatttggaatactcctagtttctctctctctctctctctctctctctctctctct 8955
QY TCGCCCCCAGCCTATGAGCATCTGAGCCCCGCCCTCTTCTTGACGAAACTTGCCCCCG 8697
Db tcgcgcccgagccta tgagcatcctgagccccgccccctctctctctgacgaaactggccccg 9015
QY ATCAGAGCAGGACCTCCCTCCGACCTCTGGAACCTCCAGAGTCCAGAGTCCAGCATCTCG 8757
Db atcagagcaggaactccctccagccctctggaaacctccagaggtccagcccatctcg 9075
QY GAGCATCCCGGAGGAATCTGACAGGGGTTAGAGTGGGTGACAGAGCTGATCTCTTT 8817
Db gagcatcccgaggaatactctgcaga-gggttagagtgaggtagacaagagcctgatctctt 9134
QY CTGTTTTGTACATAGATTTATTTTTCAGTTCACAGAAAGATGAATACATTTGTTTAAA 8877
Db cctgtttgtacatagatttttttcagttccaagaagaatgaatatactttgtlaaa 9194
QY AAAATAAAGCGCAAGTGCATTTATCTGGAAATTTGGGATGGGCGGAGTGGAG 8937
Db aaataaagcgaagtcctatctatctggaaatcgggagatggggcgaggagtgga 9254
QY CGCCCCCT 8997
Db cgccccct 9314
QY CGAGCAGAGTGGCGGAGGAGCAGTAGGCTCCGCCCGGCTTGGGGCTTCCCGCGCG 9057
Db cgagcagagtgggcggaagagcagtaggctccgcgcgcgcgcgcgcgcgcgcgcgcgcgc 9374
QY CGCCCGAGGCGGCTCCCGGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9117
Db cgcccgagggcccgctcccgcgccctctctctctctctctctctctctctctctctctct 9433
QY GCGGCCCCCGCCCCCGGCTTTCAGCCCCCGCCCCCGGCTTTCAGAGCCACGCGGCGCCGC 9177
Db gggcgcccgcccccggttccagcccccgcccccgcccccgcccccgcccccgcccccg 9493
QY 9178 CCG 9180
Db 9494 ccg 9496
|||

RESULT 5

US-09-760-475-3945/c
; Sequence 3945, Application US/09760475
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT249
; CURRENT APPLICATION NUMBER: US/09/760,475
; CURRENT FILING DATE: 2001-01-16
; Prior application data removed - consult PALM or file wrapper

NUMBER OF SEQ ID NOS: 4122
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3945
; LENGTH: 8185
; TYPE: DNA
; ORGANISM: Homo sapiens
; -09-760-475-3945

Query Match 87.9%; Score 8067.2; DB 6; Length 8185;
Best Local Similarity 99.88; Pred. No. 3.1e-150;
Matches 819; Conservative 0; Mismatches 3; Indels 10; Gaps 10;

Qy 699 GGTTAAGTGGTTCACGCCCTGTAATTTTAGCACCTTTGGGAGGCTGAGCGGGGAGATCAC 758
Db 8185 GGTTAAGTGGTTCACGCCCTGTAATTTTAGCACCTTTGGGAGGCTGAGCGGGGAGATCAC 8126
Qy 759 TTGAGGTTAGGGTTCAGACCACTGCGGCACATGGTGAACCCCATCTCTACTTAAA 818
Db 8125 TTGAGGTTAGGGTTCAGACCACTGCGGCACATGGTGAACCCCATCTCTACTTAAA 8066
Qy 819 ATACAAAAGTTTCTGGGTGTGTGGCGGACGCTATAATCCACAGTACTTGGGACTTT 878
Db 8065 ATACAAAAGTTTCTGGGTGTGTGGCGGACGCTATAATCCACAGTACTTGGGACTTT 8006
Qy 879 TTTTAAAGCGGAATCTCACTCTGTTGCCAGGCTGGAGTGCAGTGCACAGATTTCTGGCT 938
Db 8005 TTTTAAAGCGGAATCTCACTCTGTTGCCAGGCTGGAGTGCAGTGCACAGATTTCTGGCT 7946
Qy 939 CACTGAAGCCTCCGCCCTCCAGGTTCAAGGGGATTCGCCGGCTCAGCCTCCCAAGTAG 998
Db 7945 CACTGAAGCCTCCGCCCTCCAGGTTCAA -GGGATTCGCCGGCTCAGCCTCCCAAGTAG 7887
Qy 999 CTGGGAATCCCTGTCTCTGCAAAAAAATAAAAAAATAAATAAATAAATAAATAAATA 1058
Db 7886 CTGGGAATCCCTGTCTCTGCAAAAAAATAAAAAAATAAATAAATAAATAAATAAATA 7827
Qy 1059 TATATATGT 1118
Db 7826 TATATATGT 7767
Qy 1119 ACATACACAAAAATTAGCGGGAGTGTGGCGCAGCCTGTGATCACAGTACTCGGGA 1178
Db 7766 ACATACACAAAAATTAGCGGGAGTGTGGCGCAGCCTGTGATCACAGTACTCGGGA 7707
Qy 1179 GCCTGAGCAGCAGAAATCGCTTGAGCCCGTGAAGTCGAGGCTGAGCGCCAGATCGA 1238
Db 7706 GCCTGAGCAGCAGAAATCGCTTGAGCCCGTGAAGTCGAGGCTGAGCGCCAGATCGA 7647
Qy 1239 GCCACTGATTCACGCTGGCGGAAAGAGAACCGTGTCTCAAAACAACAACAAAA 1298
Db 7646 GCCACTGATTCACGCTGGCGGAAAGAGAACCGTGTCTCAAAACAACAACAAAA 7587
Qy 1299 GCTACTCTTAGCAGCTGTAGAGTATCTCGGGCGGAGTGGGAAACGAGTGTGACACA 1358
Db 7586 GCTACTCTTAGCAGCTGTAGAGTATCTCGGGCGGAGTGGGAAACGAGTGTGACACA 7527
Qy 1359 CAGAGTAGGCACTTTATATGTTAAACAGACACTGATACCCAGCTAAAGCGGCTGAACACA 1418
Db 7526 CAGAGTAGGCACTTTATATGTTAAACAGACACTGATACCCAGCTAAAGCGGCTGAACACA 7467
Qy 1419 TTTACTCTCTGGCAGTGTAAAAAGTATCTGTTTTCATATATTTTATTTTAAATTTT 1478
Db 7466 TTTACTCTCTGGCAGTGTAAAAAGTATCTGTTTTCATATATTTTATTTTAAATTTT 7407
Qy 1479 TTTGGATCAAGCAACCTGATCTTTTCCCTCATAACTTGGCGGACCGCTGACACAA 1538
Db 7406 TTTGGATCAAGCAACCTGATCTTTTCCCTCATAACTTGGCGGACCGCTGACACAA 7347
Qy 1539 AACGGCAGAGCTCGCGGACCTCCACCCCGAGTCTGCAGGTAGTGCCTCCCGGACTACA 1598
Db 7346 AACGGCAGAGCTCGCGGACCTCCACCCCGAGTCTGCAGGTAGTGCCTCCCGGACTACA 7287

Qy 1599 TTTTCCAGAAAGGCACCTCGGGCACAGCTTCTCGCTGTCTCGGCTGAAGGAAGGCCCAAT 1658
Db 7286 TTTTCCAGAAAGGCACCTCGGGCACAGCTTCTCGCTGTCTCGGCTGAAGGAAGGCCCAAT 7227
Qy 1659 CCCTGAGTATCTCGGGAAGGAGGTGTCGGGAGCGCGGACCTAGAGATCCAGAAAGCCAC 1718
Db 7226 CCCTGAGTATCTCGGGAAGGAGGTGTCGGGAGCGCGGACCTAGAGATCCAGAAAGCCAC 7167
Qy 1719 AGGCACGCGCGCGCGCCGCTATTTTCCAGGCTCTCGGGGAGCGGCTGGGCCACAGAC 1778
Db 7166 AGGCACGCGCGCGCGCCGCTATTTTCCAGGCTCTCGGGGAGCGGCTGGGCCACAGAC 7107
Qy 1779 TCCACTTCCCGGGGGTAGTGCACCTAGGGGGGAGTTCATATGCCAGCAGGCTCCG 1838
Db 7106 TCCACTTCCCGGGGGTAGTGCACCTAGGGGGGAGTTCATATGCCAGCAGGCTCCG 7047
Qy 1839 GGGGCGCTGCGCGCGCTAGTGTGGATGTAACGCGGAGTGGCGAGGGGAGCC 1898
Db 7046 GGGGCGCTGCGCGCGCTAGTGTGGATGTAACGCGGAGTGGCGAGGGGAGCC 6987
Qy 1899 GAGGCCAGGACTCTCTCTTGGGGTTTGGGGCTTGACCTGGGTGGCTTTCTTGGACAGACT 1958
Db 6986 GAGGCCAGGACTCTCTCTTGGGGTTTGGGGCTTGACCTGGGTGGCTTTCTTGGACAGACT 6927
Qy 1959 TTAGACCCCGGGGGGACAGTCTGAGAGGGGGGGGGGGCCATTTGGGGCTTCTCAT 2018
Db 6926 TTAGAG -CCCGGGGGACAGTCTGAGAGGGGGGGGGGGCCATTTGGGGCTTCTCAT 6868
Qy 2019 TGGGGTCTTGGGGCGCACCCCTACCGGCTGCCGGAATTTGTGGGGACAAAA 2078
Db 6867 TGGGGTCTTGGGGCGCACCCCTACCGGCTGCCGGAATTTGTGGGGACAAAA 6808
Qy 2079 AGGCTCTGAGTCTCGGCTGAGGGTCTCACCGACAAAAAG -AGGGGAAGCTGGTGAGCAG 2137
Db 6807 AGGCTCTGAGTCTCGGCTGAGGGTCTCACCGACAAAAAGAGGCGGTGAGCAG 6748
Qy 2138 AGGCTGAAGAGGTGGGGAAGCAGGAGGAGTGTGCTGTGCGGGGGGGTGGAACTTGA 2197
Db 6747 AGGCTGAAGAGGTGGGGAAGCAGGAGGAGTGTGCTGTGCGGGGGGGTGGAACTTGA 6688
Qy 2198 GCGGACCTTGGGAGGAGCTCCCGGGCGGAACCTGCCGACCTTCCCTCCCGGGCTTGC 2257
Db 6687 GCGGACCTTGGGAGGAGGCTCCCGGGCGGAACCTGCCGACCTTCCCTCCCGGGCTTGC 6628
Qy 2258 CTTGACGCGCGCGCCACCGCGCGCTGAGGGCTGAAAGCGCGCCCGCCAGAACGCTG 2317
Db 6627 CTTGACGCGCGCGCCACCGCGCGCTGAGGGCTGAAAGCGCGCCCGCCAGAACGCTG 6568
Qy 2318 TTTCCGCTGCGCTCCATCGACGAGTGGTGGGCTGTTTGTGCTGCGGAGCTGGGGCGAGAG 2377
Db 6567 TTTCCGCTGCGCTCCATCGACGAGTGGTGGGCTGTTTGTGCTGCGGAGCTGGGGCGAGAG 6508
Qy 2378 GAGCGGACCTTGGTCTCTTCTTCTGCTGGGCTTCTGAGGAGCATTTTCTGGCTGTC 2437
Db 6507 GAGCGGACCTTGGTCTCTTCTTCTGCTGGGCTTCTGAGGAGCATTTTCTGGCTGTC 6448
Qy 2438 AACCGCTCATCTTCCACAAAGTTTCCCGAGCTCACTTCCAGCCAGCCCGCCCGCCAG 2497
Db 6447 AACCGCTCATCTTCCACAAAGTTTCCCGAGCTCACTTCCAGCCAGCCCGCCCGCCAG 6388
Qy 2498 CCGCTTGGGGCTTCACTACTTTCGCTGGCGGCTGCTATCATATCGCGGCTCTAT 2557
Db 6387 CCGCTTGGGGCTTCACTACTTTCGCTGGCGGCTGCTATCATATCGCGGCTCTAT 6328
Qy 2558 GCCCGCTTACCGCCAGATCCGAGGGCGCTGACCTGTCCCTCTATCTCTGAGAAAGG 2617
Db 6327 GCCCGCTTACCGCCAGATCCGAGGGCGCTGACCTGTCCCTCTATCTCTGAGAAAGG 6268
Qy 2618 GGTCTCTCCAGCGGTGAGTGGTGAAGAGGTCTCCGATGTATATGAACAGCCTCAGC 2677
Db 6267 GGTCTCTCCAGCGGTGAGTGGTGAAGAGGTCTCCGATGTATATGAACAGCCTCAGC 6208
Qy 2678 CGCTCTCTTCAAGGATCGGGGCCACATCCAGTCCCTCTTTCAGCTTTCATCACAGGTGG 2737

Db 6207 CGCTCCTACTTCAAGGATCGGGCCACATCCAGTCCCTCTTACGTTTCATCACAGGTGG 6148
QY 2738 AGCCAGTAGGTGGGAATCTTATCCATGACCCACTTCTTCAAAACCCCTCCATGTTTACA 2797
Db 6147 AGCCAGTAGGTGGGAATCTTATCCATGACCCACTTCTTCAAAACCCCTCCATGTTTACA 6088
QY 2798 GAACCCCTTTTAAAGACTGTAAGCCTTGTGAGGTTCCGGCAGGTATTTTCCCTTTTGA 2857
Db 6087 GAACCCCTTTTAAAGACTGTAAGCCTTGTGAGGTTCCGGCAGGTATTTTCCCTTTTGA 6028
QY 2858 GTTGGAAACTGAAGCCCAAGAGAGGGGAATGATATGCCAAAGTCACACAGGCATGGCA 2917
Db 6027 GTTGGAAACTGAAGCCCAAGAGAGGGGAATGATATGCCAAAGTCACACAGGCATGGCA 5968
QY 2918 GGCTCGAAGTGAAGCCTGATCACTTGGCTCCAAATCATCAACCTCACCTCTGCCCCCTC 2977
Db 5967 GGCTCGAAGTGAAGCCTGATCACTTGGCTCCAAATCATCAACCTCACCTCTGCCCCCTC 5908
QY 2978 AGCACCCCAACCTTGCACCTGAACAGCTACAGGAGTTCTAAGCATGACACAGAGGGC 3037
Db 5907 AGCACCCCAACCTTGCACCTGAACAGCTACAGGAGTTCTAAGCATGACACAGAGGGC 5848
QY 3038 GGACGAGATTTAGGGGGCAAGAGATGAATTTGGGCTGCATTTGAGGAGTTTAAACAAA 3097
Db 5847 GGACGAGATTTAGGGGGCAAGAGATGAATTTGGGCTGCATTTGAGGAGTTTAAACAAA 5788
QY 3098 ATAAATGGCTATGAAGATTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGACAGGCTCTCACTCTGCC 3157
Db 5787 ATAAATGGCTATGAAGATTTTTTTTTTTTTTTTTTTTTTTTTTTTGTGACAGGCTCTCACTCTGCC 5728
QY 3158 CCAGGCTGGAGTGCAGTGTGATCATGGCTCACCTGCAGCCCTCAGCTCTCCCTGGGCTCA 3217
Db 5727 CCAGGCTGGAGTGCAGTGTGATCATGGCTCACCTGCAGCCCTCAGCTCTCCCTGGGCTCA 5668
QY 3218 GAGATCCTCAACCTCAGGCTCCTGAGTACCTGAGAGTACAGGCATGCACCGTGGTGTG 3277
Db 5667 GAGATCCTCAACCTCAGGCTCCTGAGTACCTGAGAGTACAGGCATGCACCGTGGTGTG 5608
QY 3278 GTTAATTTTTTGTATTTTTTTTTTTTGTACAGATGGTCTCACTATGTGGCCAGACTGGTCT 3337
Db 5607 GTTAATTTTTTGTATTTTTTTTTTTTGTAGAGATGGTCTCACTATGTGGCCAGACTGGTCT 5548
QY 3338 TGAACCTTTGGGCTCAAGTGCAGTGCAGCTCCGCTCAGTCTCCCAATGCTGGGATTACAGGT 3397
Db 5547 TGAACCTTTGGGCTCAAGTGCAGTGCAGCTCCGCTCAGTCTCCCAATGCTGGGATTACAGGT 5488
QY 3398 GTGAGCCACCGCAACTGGTGGCTATGAAAATTTTTTTTTTTTTTTTTCAGAGCGGCTCTAC 3457
Db 5487 GTGAGCCACCGCAACTGGTGGCTATGAAAATTTTTTTTTTTTTTTTTCAGAGCGGCTCTAC 5428
QY 3458 TCTGTCCGCCAGGCTGGAGTGCAGTGGTGAATCTCGGCTCACTGCAAGCTCTCGCTCCT 3517
Db 5427 TCTGTCCGCCAGGCTGGAGTGCAGTGGTGAATCTCGGCTCACTGCAAGCTCTCGCTCCT 5368
QY 3518 GCTTTTCATGCCATTCCTCGCTCCTCGCTCAGCCCTCCTGAGTACCTGGAGTACAGGAG 3577
Db 5367 GCTTTTCATGCCATTCCTCGCTCCTCGCTCAGCCCTCCTGAGTACCTGGAGTACAGGAG 5308
QY 3578 CTTGCCACCATGCTCGGCTAAATTTTTTTTTTTTGGATTTTGTAGTACAGAGGTTTCCACAT 3637
Db 5307 CTTGCCACCATGCTCGGCTAAATTTTTTTTTTTTGGATTTTGTAGTACAGAGGTTTCCACAT 5248
QY 3638 GTTAGCAGGATGGTCTGCATCTCTGACCTCGTGTATCCGCCGCTTGGCCCTCCCAAG 3697
Db 5247 GTTAGCAGGATGGTCTGCATCTCTGACCTCGTGTATCCGCCGCTTGGCCCTCCCAAG 5188
QY 3698 TGCTGGGATTACAGGCTGAGGCCACCGCACCTGGTCAAAATGTTTGTAGACAGAGAGGG 3757
Db 5187 TGCTGGGATTACAGGCTGAGGCCACCGCACCTGGTCAAAATGTTTGTAGACAGAGAGGG 5128
QY 3758 GCTTGACCTCAAAAGCCTTAAGAGTCAGGCTTGGAAAGAGCTTTGCACCAAGCCCGGTT 3817
|||||

Db 5127 GCTTGACCTCAAAAGGCTTAAGAGTCAGGCTTGCAAAAGAGCTTTGCACCAAGCTCGGTT 5068
QY 3818 GACTGCAATCCCATCTCTGCTGTCATATTTAGAGGAATCAGAGCTCTCTTCTCAGCT 3877
Db 5067 GACTGCAATCCCATCTCTGCTGTCATATTTAGAGGAATCAGAGCTCTCTTCTCAGCT 5008
QY 3878 TAGCAGAAAAGAGTCAGAGATAAATAGGGTTATTTGTTGGTGGGTATATAGCCAGAG 3937
Db 5007 TACGAGAAAAGAGTCAGAGATAAATAGGGTTATTTGTTGGTGGGTATATAGCCAGAG 4948
QY 3938 AGTGTGGCCAGCGCTCTGTTTGGCCATTCCTGTTTAACTAGTAGTAAAGTGCAGTAAAT 3997
Db 4947 AGTGTGGCCAGCGCTCTGTTTGGCCATTCCTGTTTAACTAGTAGTAAAGTGCAGTAAAT 4888
QY 3998 GGAATCCCTAAATCCATAGATAAATAGAGTTCAGAGAAAGAGAGTGGGCCAA 4057
Db 4887 GGAATCCCTAAATCCATAGATAAATAGAGTTCAGAGAAAGAGAGTGGGCCAA 4828
QY 4058 AGCTGGGTCAGCTACAGGATATCCAGAAAGGTATCTTGTGGACATAGAGGTTGTAAC 4117
Db 4827 AGCTGGGTCAGCTACAGGATATCCAGAAAGGTATCTTGTGGACATAGAGGTTGTAAC 4768
QY 4118 AGGAGAGAGTCTTTGAACACGCTGGAGGGAAGGATGGAGGATAGTGGCGAGGAAAT 4177
Db 4767 AGGAGAGAGTCTTTGAACACGCTGGAGGGAAGGATGGAGGATAGTGGCGAGGAAAT 4708
QY 4178 CTGAGGTTGGGTCACAGGCTTGAAGGGAGTGGAGGGAGTGGGCCATCACTACCTG 4237
Db 4707 CTGAGGTTGGGTCACAGGCTTGAAGGGAGTGGAGGGAGTGGGCCATCACTACCTG 4648
QY 4238 G-CCCCCTTTCCCATGTTAAAGCACAGAGAGCCCTCTTTTACCTACCTCCCTTCCACAG 4296
Db 4647 GCCCCCTTTCCCATGTTAAAGCACAGAGAGCCCTCTTTTACCTACCTCCCTTCCACAG 4588
QY 4297 GCACAAATTTGGACAGCTCCGCTGTCGCTGGCTTGGGCTTGGGCTTGGGCTTGG 4356
Db 4587 GCACAAATTTGGACAGCTCCGCTGTCGCTGGCTTGGGCTTGGGCTTGGGCTTGG 4528
QY 4357 GTCTCGGGATGTCACCTCGCCCTGTCGAGGATATGCTGGGTAGTGTGTTGGGCCCA 4416
Db 4527 GTCTCGGGATGTCACCTCGCCCTGTCGAGGATATGCTGGGTAGTGTGTTGGGCCCA 4468
QY 4417 ATGGGAGCAGACAGCTGAGGTGTCACCTGGCAGCAGGCAAGGCAAGGACCGCAGGGGCC 4476
Db 4467 ATGGGAGCAGACAGCTGAGGTGTCACCTGGCAGCAGGCAAGGCAAGGACCGCAGGGGCC 4408
QY 4477 AGACAGTCAATGCGGCTGGCTGAGGGGTATTTGCCCTCCCTCCCTCCCTTGTCCCT 4536
Db 4407 AGACAGTCAATGCGGCTGGCTGAGGGGTATTTGCCCTCCCTCCCTCCCTTGTCCCT 4348
QY 4537 TCATCTGTAGTAGCCCAAGCCACCCAGGAGTCCCATTTTCTTGGGCCACACCCCTTTC 4596
Db 4347 TCATCTGTAGTAGCCCAAGCCACCCAGGAGTCCCATTTTCTTGGGCCACACCCCTTTC 4288
QY 4597 TTCCCATCACCCCAATAGGAAGGAAGACAGAGAGCCCTTTTCTGGCTGTCAAT 4656
Db 4287 TTCCCATCACCCCAATAGGAAGGAAGACAGAGAGCCCTTTTCTGGCTGTCAAT 4228
QY 4657 CCCTGAAGCAGGACAGGCTGGGCCATCATGAGACATAATGATCTATCCCCCTTGA 4716
Db 4227 CCCTGAAGCAGGACAGGCTGGGCCATCATGAGACATAATGATCTATCCCCCTTGA 4168
QY 4717 GCTGGCTGTACCTGAAAGGATCATACATGCTGTGACCCCAAGATGAGGTGGCTTCA 4776
Db 4167 GCTGGCTGTACCTGAAAGGATCATACATGCTGTGACCCCAAGATGAGGTGGCTTCA 4108
QY 4777 TGTGTGTGCCATCAACCCCTTCCATTTGACCTGCACACCCGACTCGCTGGAGCTTCTGCAGC 4836
Db 4107 TGTGTGTGCCATCAACCCCTTCCATTTGACCTGCACACCCGACTCGCTGGAGCTTCTGCAGC 4048
QY 4837 TGACAGAGGTGAGGCTGAGCCCAATGGGGCAGGACTGGGCTAGGCCAGACTTGTGACTTGTCT 4896
Db 4047 TGACAGAGGTGAGGCTGAGCCCAATGGGGCAGGACTGGGCTAGGCCAGACTTGTGACTTGTCT 3988
|||||

Qy	4897	GTGGGACCTTGGCAGGGGACATTTCCCTTCTGAGCTTCAGCTTCCCCTCTCTGGAAAAA	4956
Db	3987	GTGGGACCTTGGCAGGGGACATTTCCCTTCTGAGCTTCAGCTTCCCCTCTCTGGAAAAA	3928
		TTGGGTTAGTAATTTCCCTGGCTCGCCCTTTTCCAGGGCTCTTGGGAGAGTAGAATTTGAGATG	5016
Db	3927	TGGGTTAGTAATTTCTGGCTCGCCCTTTCCACAGGCTCTTGGGAGAGTACGATTTGAGATG	3868
Qy	5017	TGAAATTTGCTTTGACTTCCATTAAGAGGCTGTGTCCACAGAAATTTTGGCCCTTCCACATGGTG	5076
Db	3867	TGAAATTTGCTTTGACTTCCATTAAGAGGCTGTGTCCACAGAAATTTTGGCCCTTCCACATGGTG	3808
Qy	5077	GGTGGTCCCTTGTGGTCTGACCCCACTCTGCCCGATAGGCTAAGGACCCGTTCTCCT	5136
Db	3807	GGTGGTCCCTTGTGGTCTGACCCCACTCTGCCCGATAGGCTAAGGACCCGTTCTCCT	3748
Qy	5137	CCCTGTTCCGTTGGCTCATAACTCTCTCTTGGGCTCTAGAAAGTGTCTGGGTGTGCTA	5196
Db	3747	CCCTGTTCCGTTGGCTCATAACTCTCTCTTGGGCTCTAGAAAGTGTCTGGGTGTGCTA	3688
Qy	5197	TGACCTGGGACATCTGGAAGTCAGTAGAGGAAGTGGCCAGGCTGGGCTGTGAGGC	5256
Db	3687	TGACCTGGGACATCTGGAAGTCAGTAGAGGAAGTGGCCAGGCTGGGCTGTGAGGC	3628
Qy	5257	CGGGGGCTGGGTGGCAGCTCAATATGATCTTCCTAGGTACCCATGGCCTTAGGG	5316
Db	3627	CGGGGGCTGGGTGGCAGCTCAATATGATCTTCCTAGGTACCCATGGCCTTAGGG	3568
Qy	5317	AACTTGGCAGATCTAGAGGAGCTGGAGGCCACCCCTGGCCGGCCAGACCCACTCACCCCTC	5376
Db	3567	AACTTGGCAGATCTAGAGGAGCTGGAGGCCACCCCTGGCCGGCCAGACCCACTCACCCCTC	3508
Qy	5377	TACCAAGGTTGGGGCATCTAAGAGAGGTTGCAGAGGGAGACCCCTAACAGTGGCTGAGG	5436
Db	3507	TACCAAGGTTGGGGCATCTAAGAGAGGTTGCAGAGGGAGACCCCTAACAGTGGCTGAGG	3448
Qy	5437	CAGGGGCCCTCATCTGGGCAGATGAGAAGAACTTTGTGTGTGGGGGTATCGCCCAT	5496
Db	3447	CAGGGGCCCTCATCTGGGCAGATGAGAAGAACTTTGTGTGTGGGGGTATCGCCCAT	3388
Qy	5497	CCAGTCTACATTTGTGTCAACTGTGTGCAGAAATCAGTTTCACTAGGCGTGTCTGAGGGGT	5556
Db	3387	CCAGTCTACATTTGTGTCAACTGTGTGCAGAAATCAGTTTCACTAGGCGTGTCTGAGGGGT	3328
Qy	5557	GTCCAGGGTTCCCAAGCCTGGAGTGGCAGGGGTGCATTTGTCCCTCAGCCCTGCCTT	5616
Db	3327	GTCCAGGGTTCCCAAGCCTGGAGTGGCAGGGGTGCATTTGTCCCTCAGCCCTGCCTT	3268
Qy	5617	TTCTGCCACTGTCTTACTGTCTTCTCGAGTATACAGAGTCAAAATGTGTAGGAGCAC	5676
Db	3267	TTCTGCCACTGTCTTACTGTCTTCTCGAGTATACAGAGTCAAAATGTGTAGGAGCAC	3208
Qy	5677	TGAGAGGGGGTGTTCACATTCGTGGGTGTAGTGGGAGGAGGCCATTTGGGCTGGGCTT	5736
Db	3207	TGAGAGGGGGTGTTCACATTCGTGGGTGTAGTGGGAGGAGGCCATTTGGGCTGGGCTT	3148
Qy	5737	GAAAGTCTTTGGTGTATGTAGAAGAGTGTCTTGAAAAAGAAAGGGGCCCTGAGCTCGGAG	5796
Db	3147	GAAAGTCTTTGGTGTATGTAGAAGAGTGTCTTGAAAAAGAAAGGGGCCCTGAGCTCGGAG	3088
Qy	5797	GGCAGGCCACCCCTTGAGTCTGCCCAAGGCTCAGCCAGCATGCTCTGTAGACCCAGGG	5856
Db	3087	GGCAGGCCACCCCTTGAGTCTGCCCAAGGCTCAGCCAGCATGCTCTGTAGACCCAGGG	3028
Qy	5857	AGGAGCACAGGTAGAAGGCTGGCAGCATGGAGTGGGAGTGGAGATGGAGAGGACTC	5916
Db	3027	AGGAGCACAGGTAGAAGGCTGGCAGCATGGAGTGGGAGTGGAGATGGAGAGGACTC	2968
Qy	5917	CTGGGATCTTCTGTGGCCCTCTTGGGTGTGCCCCCTGGTGGGGATTTGTGCCAGCAGG	5976
Db	2967	CCTGGATCTTCTGTGGCCCTCTTGGGTGTGCCCCCTGGTGGGGATTTGTGCCAGCAGG	2908

Qy	5977	GCAGCTGGGGCTGCCCTCCCTGAGGATCCTCTGCCTCACCCTCCCATCCAGGGCATTCGCTCA	6036
Db	2907	GCAGCTGGGGCTGCCCTCCCTGAGGATCCTCTGCCTCACCCTCCCATCCAGGGCATTCGCTCA	2848
Qy	6037	GCCAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACCTGGCTGGGTACACAC	6096
Db	2847	GCCAAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACCTGGCTGGGTACACAC	2788
Qy	6097	TGTCGCAACCCAAATGTGCGGGAAGCCCTGAGGCCCTGGGGGACACGGGAACTGTCATC	6156
Db	2787	TGTCGCAACCCAAATGTGCGGGAAGCCCTGAGGCCCTGGGGGACACGGGACACTGTTCATC	2728
Qy	6157	CAGGAGTGAGGATCCCCCTACTAGGGCTGCAGAGCTGTCTTCTTCTCCCTCCCATCAGTT	6216
Db	2727	CAGGAGTGAGGATCCCCCTACTAGGGCTGCAGAGCTGTCTTCTTCTCCCTCCCATCAGTT	2668
Qy	6217	TCCAACCAACCCTCGTCCAGGACTGAGGCCTGGCTCCCAAGCCCAATCCCTTTCCATCCA	6276
Db	2667	TCCAACCAACCCTCGTCCAGGACTGAGGCCTGGCTCCCAAGCCCAATCCCTTTCCATCCA	2608
Qy	6277	GTCCCTAGGACGAAGGCCACCATTTACCCAGGAGTGGGACCTGATTTAAGTGTCTACA	6336
Db	2607	GTCCCTAGGACGAAGGCCACCATTTACCCAGGAGTGGGACCTGATTTAAGTGTCTACA	2548
Qy	6337	TCCTTTGCCCTCCCTCCCTCCTCTCTTAATTTTCTTCTCAGAAACAGTCTCAAACTCC	6396
Db	2547	TCCTTTGCCCTCCCTCCCTCCTCTCTTAATTTTCTTCTCAGAAACAGTCTCAAACTCC	2488
Qy	6397	AATGTTTAAACACCATCATCCAGCAGTGGGACTTCCACCCTCGCGCCCATGCGCCCCCTCC	6456
Db	2487	AATGTTTAAACACCATCATCCAGCAGTGGGACTTCCACCCTCGCGCCCATGCGCCCCCTCC	2428
Qy	6457	TCATTCCTGCTTCTCTCTGGCTGACCCAGACAGCATATTTTCAGCTGAGGACCC	6516
Db	2427	TCATTCCTGCTTCTCTCTGGCTGACCCAGACAGCATATTTTCAGCTGAGGACCC	2368
Qy	6517	ACCTACTCCCCAGCCCTCGGGGCTCATCCCGCGCAGGTCCCTGGGGCTACCCCGCA	6576
Db	2367	ACCTACTCCCCAGCCCTCGGGGCTCATCCCGCGCAGGTCCCTGGGGCTACCCCGCA	2308
Qy	6577	TGGTGAGACCCCTTCAGACCTACAGAGACCCACTGCTCTCAGAGCTACAACTACTGCC	6636
Db	2307	TGGTGAGACCCCTTCAGACCTACAGAGACCCACTGCTCTCAGAGCTACAACTACTGCC	2248
Qy	6637	GGGAAGACGAGGAGATCTACAAGGAGTTCTTTGAAGTAGCCAATGATGTATCCCAACC	6696
Db	2247	GGGAAGACGAGGAGATCTACAAGGAGTTCTTTGAAGTAGCCAATGATGTATCCCAACC	2188
Qy	6697	TGCTGAAGGAGGACAGCACTTTGTGAGGCGGCGAGGAGCGCGGGGAGCAAGCC	6756
Db	2187	TGCTGAAGGAGGACAGCACTTTGTGAGGCGGCGAGGAGCGCGGGGAGCAAGCC	2128
Qy	6757	AGGTGAAGGCTGAGCTCCAGCCTGTCTCAGAGCTCCCACTCGGACAGGCTCCCTTCC	6816
Db	2127	AGGTGAAGGCTGAGCTCCAGCCTGTCTCAGAGCTCCCACTCGGACAGGCTCCCTTCC	2068
Qy	6817	ACAGGGCATGGGGCTGCATGTACGGATTAGGGATGGCAGGAGAACTTGGCCCTGAG	6876
Db	2067	ACA - GGCCATGGGGCTGCATGTACGGATTAGGGATGGCAGGAGAACTTGGCCCTGAG	2009
Qy	6877	CAGACAGTATGTCCTTTTTCATTAACCTGAGTCTGGGCGCCAGCTTGACGGGACTG	6936
Db	2008	CAGACAGTATGTCCTTTTTCATTAACCTGAGTCTGGGCGCCAGCTTGACGGGACTG	1949
Qy	6937	AAGGTATTTTAGAGTTTCTACCTGTGCCCTCAGTTTTCATGGCCAGACTCCCTCCCTCA	6996
Db	1948	AAGGTATTTTAGAGTTTCTACCTGTGCCCTCAGTTTTCATGGCCAGACTCCCTCCCTCA	1889
Qy	6997	GCTGAGGGTGAGGTAGGATGTAGTACGTCTGCGCTATGGATTGGCTTTATAAAGGAAA	7056
Db	1888	GCTGAGGGTGAGGTAGGATGTAGTACGTCTGCGCTATGGATTGGCTTTATAAAGGAAA	1829
Qy	7057	GAGGTCTTAAGAAATGTTCCCAACCATGCTTTACCTTTTCTTGAGGCCAGGGGTCTTTGGCT	7116

Db 1828 GAGGTTCTAAGAAATGTTCCCAACCTATGCTTACCTTTCTCGAGCCAGGGGTCTTTGCCT 1769
QY 7117 AGTGGGGGCGCTGGCTGTGCGCTCTGCTAAGGGGTGAGTAAAGACTGATCTGTGCC 7176
Db 1768 AGTGGGGGCGCTGGCTGTGCGCTCTGCTAAGGGGTGAGTAAAGACTGATCTGTGCC 1709
QY 7177 TCCTTTCCCTCCCTCGTCAGGGCACCCAGAGCCAAAGGTTCGCGCTCCAGGACCTGAGTG 7236
Db 1708 TCCTTTCCCTCCCTCGTCAGGGCACCCAGAGCCAAAGGTTCGCGCTCCAGGACCTGAGTG 1649
QY 7237 CTTTCGCCACCTCGCTCGGATTTACAGCGGCATCTGCAATGGAGGAGGCGAGTCCAC 7296
Db 1648 CTTTCGCCACCTCGCTCGGATTTACAGCGGCATCTGCAATGGAGGAGGCGAGTCCAC 1589
QY 7297 GCCTGTGCTGACGTGGCTGGCGCACCTTTCTGTGTCAGTCCCTAGGCGGTTTTCAGGG 7356
Db 1588 GCCTGTGCTGACGTGGCTGGCGCACCTTTCTGTGTCAGTCCCTAGGCGGTTTTCAGGG 1529
QY 7357 ACAGGTGAGGACAGCTGACAGAGGTCTGGGCATCTACAGGTGGTGACAGCAGCCAGGG 7416
Db 1528 ACAGGTGAGGACAGCTGACAGAGGTCTGGGCATCTACAGGTGGTGACAGCAGCCAGGG 1469
QY 7417 CTTGTGAGACTTTTTCGGCCAGGGCAGCATCTGCCATTCCTTCGGTGGCGATGGGA 7476
Db 1468 CTTGTGAGACTTTTTCGGCCAGGGCAGCATCTGCCATTCCTTCGGTGGCGATGGGA 1409
QY 7477 CTGAGACCCCTCGGTGGATGGATGGCCAGAGAGGTCTGGGCATCTACAGGTGGTGACAGCAGCCAGGG 7536
Db 1408 CTGAGACCCCTCGGTGGATGGATGGCCAGAGAGGTCTGGGCATCTACAGGTGGTGACAGCAGCCAGGG 1349
QY 7537 CCGGCAACCTTGCTCTACCTTGCTCTCCCTACTGCGCCAGGTGCGCGAGAAGTGGCA 7596
Db 1348 CCGGCAACCTTGCTCTACCTTGCTCTCCCTACTGCGCCAGGTGCGCGAGAAGTGGCA 1289
QY 7597 TAGTGAGCCGAGAGCGGAGCGCGCGCGAGGCGCGAGAGCGGTGGGGGAGGAGCGCGGG 7656
Db 1288 TAGTGAGCCGAGAGCGGAGCGCGCGCGAGGCGCGAGAGCGGTGGGGGAGGAGCGCGGG 1229
QY 7657 AAGCGCGGCGCGCGCGCGCGCGAGGTGCAAGGCGAGAGGAGCGCGCGCGCGCGAAGA 7716
Db 1228 AAGCGCGGCGCGCGCGCGCGCGAGGTGCAAGGCGAGAGGAGCGCGCGCGCGCGAAGA 1169
QY 7717 AGCCAGACTTGAAGAGGCTGGGCGACCGCGAGGTGCAAGTGTGAGGACCGCGCGCGGA 7776
Db 1168 AGCCAGACTTGAAGAGGCTGGGCGACCGCGAGGTGCAAGTGTGAGGACCGCGCGCGGA 1109
QY 7777 AGCCTCTGGGACTGCTGCTGGCGACAGCGCGGAGCGCGTGAAGTGGCAGCGGCTCAGG 7836
Db 1108 AGCCTCTGGGACTGCTGCTGGCGACAGCGCGGAGCGCGTGAAGTGGCAGCGGCTCAGG 1049
QY 7837 TGCAGCAGCCGCGAGCATCACCACCGCGGAGGTCCAGTGTCACTTTCCAGAGTGAGA 7896
Db 1048 TGCAGCAGCCGCGAGCATCACCACCGCGGAGGTCCAGTGTCACTTTCCAGAGTGAGA 989
QY 7897 AGATGAAGGCGATGAAGGAGTCTGCTGGTGGCCACCAAGATCAACTCGAGGCGCATCAAGC 7956
Db 988 AGATGAAGGCGATGAAGGAGTCTGCTGGTGGCCACCAAGATCAACTCGAGGCGCATCAAGC 929
QY 7957 TGCACACTCAGGCGACAGTGCAGTGCAGTGAAGAGCAGAAAGTGTCCACCCCTAGTG 8016
Db 928 TGCACACTCAGGCGACAGTGCAGTGCAGTGAAGAGCAGAAAGTGTCCACCCCTAGTG 869
QY 8017 ACTACACTCTGCTTTCTTCAAGCGCGACGCGCAAGGCGCTCTCAACTACTGGGAGTTCG 8076
Db 868 ACTACACTCTGCTTTCTTCAAGCGCGACGCGCAAGGCGCTCTCAACTACTGGGAGTTCG 809
QY 8077 GACCGCTTGTGGGACCCAGGCTCCG-CCTTAGTCCCGCAACTCTGAGCCCATGTTCTGC 8135
Db 808 GACCGCTTGTGGGACCCAGGCTCCGCGCTTAGTCCCGCAACTCTGAGCCCATGTTCTGC 749
QY 8136 CCCCAGCCAAAGGGACAGGCGCTACCTCTACCCAAACCTTAGTTCGCGGTCCGAGT 8195
|||||

Db 748 CCCAGCCCCAAAGGGGACAGGCGCTCACCTACCCAAAACCTAGGTTCCCGGTCCCGAGT 689
QY 8196 ACAGTCTGTATCAAAACCCACAGATTTCTCCAGCTCAGAAACCCAGGGCTCTTGCCCACTCG 8255
Db 688 ACAGTCTGTATCAAAACCCACAGATTTCTCCAGCTCAGAAACCCAGGGCTCTTGCCCACTCG 629
QY 8256 TTAGAATATAGGTCCTCTTCTCCAGAGATCCAGCGGCCAATGAAACCTCACGCTGGT 8315
Db 628 TTAGAATATAGGTCCTCTTCTCCAGAGATCCAGCGGCCAATGAAACCTCACGCTGGT 569
QY 8316 CTTAATTACAGCTCTTTAAAGGCCAGCGCTTAGAAACCCAAAGCTCTCTTCGGAACCG 8375
Db 568 CTTAATTACAGCTCTTTAAAGGCCAGCGCTTAGAAACCCAAAGCTCTCTTCGGAACCG 509
QY 8376 TCACCTAGAGCCAGACCAACGTTACTCAGGCTCTCCAGGCTCTCCAGGCTTGTAGGAGCTGAGGTTT 8435
Db 508 TCACCTAGAGCCAGACCAACGTTACTCAGGCTCTCCAGGCTCTCCAGGCTTGTAGGAGCTGAGGTTT 449
QY 8436 CACCTTTAACCCAAAGGGGACAGGTCCTCCAGCTCCAGCGCGGGGAGCTTAGGACCACTCA 8495
Db 448 CACCTTTAACCCAA-GGAGCACAGGTCCCACTCCAGCCC-GGAGCCTAGGACCACTCA 391
QY 8496 GCCCTAGGAGTATATTTCCGCACTTCAGAAATTCATATCTTGGGAATCCAAAGCTCCCTG 8555
Db 390 GCCCTAGGAGTATATTTCCGCACTTCAGAAATTCATATCTTGGGAATCCAAAGCTCCCTG 331
QY 8556 CCCCATAACTTCAGTCTCTCCAGATTTCCAGATTTGGAAATCCTAGTTCTCTCTTCGTA 8615
Db 330 CCCCATAACTTCAGTCTCTGCTGC-TCCAGAAATTTGGAAATCCTAGTTCTCTCTTCGTA 272
QY 8616 TCCGAGTCTGGGACACAAACCTCCGCCCCAGCCTATGAGCATCTGAGCCCCGCCCTC 8675
Db 271 TCCGAGTCTGGGACACAAACCTCCGCCCCAGCCTATGAGCATCTGAGCCCCGCCCTC 212
QY 8676 TTCTGACGAACTGGCCCCGGATCAGAGCAGGACCTCCCTTCGACCTCTGGGAACCT 8735
Db 211 TTCTGACGAACTGGCCCCGGATCAGAGCAGGACCTCCCTTCGACCTCTGGGAACCT 152
QY 8736 CCCAGAGTCCAGCCATCTCGGAGCATCCCGGAGGAAATCTGCAGAGGGGTTAGGAGTG 8795
Db 151 CCCAGAGTCCAGCCATCTCGGAGCATCCCGGAGGAAATCTGCAGA-GGTTAGGAGTG 93
QY 8796 GGTGACAAGAGCCTGATCTCTCTCTCTCTCTCTCTGTTGTACATAGATTTATTTTCAGTTCCAAAGAA 8855
Db 92 GGTGACAAGAGCCTGATCTCTCTCTCTCTCTCTCTGTTGTACATAGATTTATTTTCAGTTCCAAAGAA 33
QY 8856 AGATGATACATTTGTTAAAAAATAATAA 8887
Db 32 AGATGATACATTTGTTAAAAAATAATAA 1

RESULT 6
PCT-US01-14827-388
; Sequence 388, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 388
; LENGTH: 2770
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2604)..(1804)
; OTHER INFORMATION: 100% homologous to Homo sapiens C33 antigen-type III integral
; OTHER INFORMATION: membrane protein, accession number S48196, Smith-Waterman Score=

OTHER INFORMATION: 1413.
PCT-US01-14827-388

Query Match 15.58; Score 1425.2; DB 1; Length 2770;
Best Local Similarity 99.1%; Pred. No. 7.6e-22;
Matches 1433; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy	2288	ATGGGGCTGAAGCGCGCCAGACAGCGCTGTCCCGCTGCGCTCCATCGACGAGCGTG	2347
Db	1	atg999ctgaag99ccgcccagaagcgtgtcccgctg9cgctccatcgacgagctg9gt	60
Qy	2348	CGCTGTGTGTCGCGAGCTGGCGCGAGAGCGGACCTGCTGCTCTTTCCTTGGTG	2407
Db	61	c9cctgttctgctcgagctg9ccgagagagcggagacctggtgctcttctctg9gt	120
Qy	2408	CTGGGCTTGTGGAGCATTTTCTGGCTGTCAACCGCGCTATCCCTACGAGCTTCCGAG	2467
Db	121	ct99gcttctg9gagcattttctgctgtcaacccgctcatccctaccacggttccc9ag	180
Qy	2468	CTCACTTCCAGCCAGCGCCCGCCCGACCGCTGGCGGCTTCACTTCCGCTG	2527
Db	181	ctcacccttccagccagcccgcccgcccgacccgctg9cgctccacttcttccc9tg	240
Qy	2528	GCCGACCTGTCTATCATGCGCGCCCTCTATGCCGCTTCAACCGCCGAGCGGCC	2587
Db	241	gccgacctgtctatcatcgccgctctctatgcccgttccacgcccagatccgagcgcc	300
Qy	2588	GTGACCTGTCCCTTATCTTCGAGAGGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAAG	2647
Db	301	gtcagacctgtccctctatctctcgagaggg9gtgtccccagccgtgagctg9tgaagaag	360
Qy	2648	GTCTCCGATGTATATGAACAGCGCTCAGCGCTCTACTTCAAGGATCGGCGCCACATC	2707
Db	361	gtctcgatgtcatatgaaacagcctcagccgctcctacttcaagatcg9gcccacatc	420
Qy	2708	CAGTCCCTCTTACGCTTCATCACAGGTTGGAGCCAGTAGTGGAATCTTATCCATGAC	2767
Db	421	cagtcctcttctcagcttctacaggttggagcccgtagtgggaatcttatccatgac	480
Qy	2768	CCACTTCTTCAAAACCTCCATGTTTACAGAACCTTTTAAAGACTGTAGCGCTTGA	2827
Db	481	ccactcttcaaaacctccatg9ttacagaaacctttaaagactgtgaagcctt9ga	540
Qy	2828	GGTTCGGCAGGTGTATTTCTCTTTCAGTTGGGAACCTGAAGCCAGAGGGGAAA	2887
Db	541	g9ttcggcaggtgttatcttctcttctgagttgggaactgaagccagagaggggaaa	600
Qy	2888	TGATATGCCAAAGTCACACAGCGGATGGCAGGGCTGGAGTGAAGCTGATCTGGCT	2947
Db	601	tgatgcccgaagtcaacacggtcatg9cagggctggaagtgaagcctgatcacttgct	660
Qy	2948	CCAAATCATCAACCTCACTCTGCGCCCTCAGCAGCCGCCCTTGCCACTGAACAGCTA	3007
Db	661	ccaaatcataacctcaacctctgccccctcagccccccaccttgcactgaaacagcta	720
Qy	3008	CAGGAGTTCTAAGCATGACACAGAGGGCGGACAGATTAGGGCGCAAGAGATGAA	3067
Db	721	caggagttcctaagcatgagacacagagggcg9cagcagatgttagggggcagaagatgaa	780
Qy	3068	ATGGGCTGCATTTGAGCGCAGTTAAACAAATAATGGCTATGAAGATTTTTTTTTT	3127
Db	781	att99gctgtcttggagcagtttaacaaaaataatggtctgaagatttttttttt	840
Qy	3128	TTTTTTTTCAGACAGGGCTCACTCTGTCCCGCAGGCTGGAGTGGTGTGATCATGG	3187
Db	841	ttttttttgagacaggggtctcactctgttcccccaaggtgagtgagtgatcatg9	900
Qy	3188	CTCACTGACGCGCTCAGTCTCCCTGGGCTCAGAGATTCCTCCAACTCAGCCTCCCTGAGTAG	3247
Db	901	ctcactgagcctcagtcctcctg9gtcagagatccctccaaacctcagccctcctgagtag	960
Qy	3248	CTGAGAGTACAGGCATGACCGTGGTGTGCTTAATTTTTTTTGTATTTTTTTGTAGAGAT	3307

Db	961	ctgagagtcagggcatgcaccgtggtgctgtaacttaatttttttttttttttttttttt	1020
Qy	3308	GGTCTCTCACTATGTGCCCCAGACTGGTCTTGAACCTCTTGGGCTCAAGTGATCTGCCCGC	3367
Db	1021	g9tgtccactatgtgcccagactggtcttgaactcttgggtccaagtgtatcyccgc	1080
Qy	3368	CTCAGTCTCCCAATGTCTGGATTACAGGTGTAGCCACCGCAACTGGTGGCCTATGAAA	3427
Db	1081	ctcagttcccaaatgctgggtattacagggtgagccacccagcaactggctg9ccatataaaa	1140
Qy	3428	ATTTTTTTTTTTTTCAGACGGCTCTCACTCTCTGCCCCAGGCTGGAGTGCAGTGTGTC	3487
Db	1141	atttttttttttttcagaagcgctctcaactctgtgcccag9ctg9agtgagtg9tgc	1200
Qy	3488	AATCTCGGCTCACTGCAAGCTCTGCTCTCTTTCATGCCATTTCTCTGCTCTCTGCT	3547
Db	1201	aatctcgctcaactgcaagctctgctctctcttctcatgcatctctctgctctgct	1260
Qy	3548	CAGCTCTCTAGTAGCTGGGACTACAGGAGCTGCCACATGCTGGCTAAATTTTTTTTTT	3607
Db	1261	cagctctcagtagctg9gagactacaggtgcccgcacccg9cccg9ctaatttttttt	1320
Qy	3608	GGATTTTATAGTAGACGAGGTTTTCACCATGTTAGCCAGGATGGTCTCGATCTCTGACC	3667
Db	1321	gtatttttagtagagaggggtttccacctgttagccaggtggtctcgatctctctgacc	1380
*Qy	3668	TCGNGATCCGCGCCCTTGGCTTCCCAAGTGTGGGATTTACAGCGTACAGCGGACCGCAC	3727
Db	1381	tc9tgatccaccccgctc9gctcccaaa9t9t9gg9attacag9c9t9g9ccac9c9c	1440
Qy	3728	CTGGTC 3733	
Db	1441	ccg9cc 1446	

RESULT 7

PCT-US01-14827-5297/c
; Sequence 5297, Application PC/TUS0114827
; GENERAL INFORMATION:
; APPLICANT: Hyseq, Inc
; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES
; FILE REFERENCE: 21272-104
; CURRENT APPLICATION NUMBER: PCT/US01/14827
; CURRENT FILING DATE: 2001-05-16
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 16102
; SOFTWARE: Custom
; SEQ ID NO 5297
; LENGTH: 2121
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIMILAR
; LOCATION: (2121)...(1651)
; OTHER INFORMATION: 98% homologous to Homo sapiens menin, accession number
; OTHER INFORMATION: U93237, Smith-Waterman score=783.
PCT-US01-14827-5297

Query Match 14.9%; Score 1365.2; DB 1; Length 2121;
Best Local Similarity 98.1%; Pred. No. 1.3e-20;
Matches 1413; Conservative 0; Mismatches 23; Indels 4; Gaps 3;

Qy	2288	ATGGGGCTGAAGCGCGCCAGACAGCGCTGTCCCGCTGCGCTCCATCGACGAGCGTG	2347
Db	2121	ATGGGGCTGAAGCGCGCCAGACAGCGCTGTTCGCGCTGCCCTCCATCGACGAGCGTG	2062
Qy	2348	CGCTGTGTGCTGCGGAGTGGCGGAGAGGAGCGGACCTGGTGTCTCTTTTCCCTTGGTG	2407
Db	2061	CGCTGTGTGCTGCGGAGTGGCGGAGAGGAGCGGACCTGGTGTCTCTTTTCCCTTGGTG	2002

Db	921	CAATCTGGCTCACTGCAAGCTCTCGCTCTGCTTTTCATGCCATTCCTCGCCTCGCC	862
Qy	3547	TCAGCCTCTCTGAGTAGCTGGGACTACAGAGAGCTGCCACCATGGCTGGCTAAATTTTTTTT	3606
Db	861	TCAGCCTCTGAGTAGCTGGGACTACAGAGAGCTGCCACCATGGCTGGCTAAATTTTTTTT	802
Qy	3507	TGATATTTTAGTAGAGAC--GAGGTTTCACCATGTTAGCCAGATGGTCTCGATCTCTGA	3655
Db	801	TGGATTTTGTAGTAGAGACAGGGGTTTCACCATGTTGGCCAGGCTGGTCTCGAACTCTGA	742
Qy	3666	CTTC--GTGATCGCCCGCTTGGCTCCCAAGTGCTGGGATACAGCGCTGAGGCCACC	3723
Db	741	CTCAGGTGATCCGCCCGCTTGGCTCCAGAGTGCTGGGATTTACAGCATGAGTGAGC	682
RESULT 8			
US-09-948-941-5			
; Sequence 5, Application US/09948941			
; GENERAL INFORMATION:			
; APPLICANT: VENTER, J. Craig et al.			
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED			
; TITLE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF			
; FILE REFERENCE: CL000788			
; CURRENT APPLICATION NUMBER: US/09/948,941			
; CURRENT FILING DATE: 2001-09-10			
; PRIOR APPLICATION NUMBER: 60/231,328			
; PRIOR FILING DATE:			
; NUMBER OF SEQ ID NOS: 12618			
; SOFTWARE: FastSeq for Windows Version 4.0			
; SEQ ID NO 5			
; LENGTH: 2772			
; TYPE: DNA			
; ORGANISM: Human			
US-09-948-941-5			
Query Match 14.3%; Score 1312.2; DB 5; Length 2772;			
Best Local Similarity 96.5%; Pred. No. 1.1e-19;			
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps			
Qy	7499	GGATGGCCAGACAGGGTCTCGTAGTTCCAGGCACCTGGCCGGCAACCTTGTCTCACCTT	7558
Db	1382	gagggcagtcacacgcctgtctgcagctggctggccaccctttcttgacgtccct	1441
Qy	7559	GCTCTCCCCTACTGGCCCGAGTGGCGCAGAAAGTGGCGCATAGTGAGCCGAGAGCCGAGGC	7618
Db	1442	agggcgttttgaggagcagtgctggcagaagctgcgcatagtgagccgagagcgcgagc	1501
Qy	7619	GGCCGAGGCGGAGAGCCCTGGGGCGAGAAAGCCGGAAGCCGGCGGGCGGGCCGACG	7678
Db	1502	ggccgagccgagagccgtg9ggcgagaaagcccg9gagcccg9ggggccacg	1561
Qy	7679	GCGGGAGTCCAACGACAGAGAGCCCGCCGCCCAAGAGCCAGCATCTGACAAAGGCGCT	7738
Db	1562	gcggagtcacagcagagagagcccgcccaagaagccagcactggacaagagcct	1621
Qy	7739	GGGACCGGCCAGGGTGAGTGTGAGGACCCCGGGAAGCCCTCTGGGATGTGCGTGG	7798
Db	1622	gggcacggcgaggtgcagtgcaaggacccccccggagcctcctgggactgctgcgtg	1681
Qy	7799	CACAGCCGAGGCGCTTGAAGTGGCAGACGCTCAGGTGCCACCCCGCAGCATCACC	7858
Db	1682	cacagcccgagccctg9aggg9ggcagcagcctcag9tgcacagcccgagcattacc	1741
Qy	7859	ACCBCGGAGGGTCCAGTGCCTCACTTTCACAGAGTGAAGATGAAGGCGCATGAAGAGCT	7918
Db	1742	accgcgagggccagtgctcactttccagagtgagaagatgaagg9caggaagagct	1801
Qy	7919	GCTGGTGCCACCAAGATCACTGAGCGCCATCAAGCTGCAACTCAGCGCACAGTGC	7978
Db	1802	gctggtggccacaaagatcaactcagcgcctcaagctgcaactcaccgacagtcgca	1861

Qy 7979 AGTCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACTCTGTCTTCTCAA 8038
Db 1862 agtcagatgaagaagcagaaagtgtccacccctagtgaactacactctgtcttctctcaa 1921
Qy 8039 GCGCAGCGCAAAAGCCCTCTGAACCTACTTGGGACCTTGGGGACCCAGGC 8098
Db 1922 gcgcagcgcaaaagccctctgaactactgggactctggagccgcttggggaccacagc 1981
Qy 8099 TCCGCCCTTAGTCCCAACTCTGAGCCCATGTTCTGCCCCAGCCCAAGGGACAGGCC 8158
Db 1982 tccgccttagtccccaactctgagccatgttctgccccagcccaaaagggacagcc 2041
Qy 8159 TCACCTCTACCAAAACCCCTAGTTCGCGGTCCCGAGTACAGTCTGTATCAAAACCCAGAT 8218
Db 2042 tcacctctacccaacccctaggttcccggtcccgagtacagctgtgtatcaaaaccacgat 2101
Qy 8219 TTTCTCCAGCTCAGAACCCAGGGCTCTGCCCCAGTCTGTAGATATAGGTCTCTTCTCC 8278
Db 2102 tttctccagctcagaacccagggctctgccccagtcgttagaataataggtctcttctccc 2161
Qy 8279 AGAATCCAGCGCGCCCAATGAAACCTCACGCTGGTCTCTAATTAACAGTCTTTAAAGGC 8338
Db 2162 agaattccagcgccgccaatgaaacctcacgctgggtccctaattacagctctttaaagcc 2221
Qy 8339 CCAGCCCTTAGAAACCCCAAGCTCTCTCGGAACCGCTCACCTAGAGCCAGACCAAGCTT 8398
Db 2222 ccagcccttagaaaccccaagctctctcctcggaacgctcaccttagagccagaccagctt 2281
Qy 8399 ACTCAGGGCTCTCCAGCTGTAGGAGCTGAGGTTCACCCCTTAACCCAGGGAGGACACA 8458
Db 2282 actcagggctctccagctgttaggagctgtaggttccaccttaaccccaaggagacaca 2341
Qy 8459 GGTCCCACTCTCAGCCCGGGAGGAGCTAGGACCACTCAGCCCTAGGAGTATATTTCCGCA 8518
Db 2342 ggtcccaactccagcccgaggagcctaggaccactcagcccttaggagtagtatatttcgca 2401
Qy 8519 CTTCAGAAATTCATATCTTGGGAATCCAAGCTCCCTGCCCAATAACTTCAGTCTCTGCT 8578
Db 2402 cttcagaattccatctcttgcgaatccaaagctcctcgccccaaataaacttcagtcctgct 2461
Qy 8579 TCCAGAAATTTGGAATCTAGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8638
Db 2462 tccagaatttggaaatccttagtttccctctctctctctctctctctctctctctctctct 2521
Qy 8639 CCGCCCCAGCCTATGAGCAATCTGAGCCCGCCCTCTCTCTGACGAAACTGCCCCCGGA 8698
Db 2522 ccgccccagcctatgagcactctgagcccgccctctctctctgacgaaactggccccgga 2581
Qy 8699 TCAGAGCAGGACCTCCCTTCGACCCCTCTGGGAACCTCCAGAGGTCCAGCCCATCTCGG 8758
Db 2582 tcagagcaggacctcccttcgagccctctggaacctcccgagaggtccagcccatctcgg 2641
Qy 8759 AGCATCCGGAGGAAATCTGAGAGGGGTTAGGAGTGGGTGACAGAGCCCTGATCTCTTC 8818
Db 2642 agcatcccgaggaaatctgcagagggggttaggagtgagggtgagcagagcctgactctctc 2701
Qy 8819 CTGTTTGTACATAGATTTATTTTCTAGTTCACAGAAAGATGAATACATTTTGTAAAAA 8878
Db 2702 ctgtttgtacatagattttattttctagttccagtaagagagataacattttgttaaaaa 2761
Qy 8879 AAATATAAA 8887
Db 2762 aaaaaaaaa 2770

RESULT 9

US-09-380-337-1

; Sequence 1, Application US/09380337

; GENERAL INFORMATION:

; APPLICANT: Chandrasekharappa, Sattara C.

; Guru, Siraadnanahalli C.

; Manickam, Pachaiappan

; Collins, Francis S.

Emmert-Buck, Michael R.
Debelenko, Larisa V.
Lubensky, Irina A.
Liotta, Lance A.
Agarwal, Sunita K.
Spiegel, Allen M.
TITLE OF INVENTION: MEN1, the Gene Associated With Multiple Endocrine Neoplasia Type 1, Menin Polypeptides, and Uses Thereof

NUMBER OF SEQUENCES: 28

CORRESPONDENCE ADDRESS:

ADDRESSEE: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, Eighth Floor
CITY: San Francisco
STATE: California
COUNTRY: USA

ZIP: 94111-3834

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/380,337

FILING DATE: 09-Mar-2000

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 60/040,269

FILING DATE: 05-Mar-1997

APPLICATION NUMBER: WO PCT/US98/04258

FILING DATE: 04-Mar-1998

ATTORNEY/AGENT INFORMATION:

NAME: Lockyer, Jean M.

REGISTRATION NUMBER: 44,879

REFERENCE/DOCKET NUMBER: 015280-315100US

TELECOMMUNICATION INFORMATION:

TELEPHONE: (415) 576-0200

TELEFAX: (415) 576-0300

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 2772 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: cds

LOCATION: 111..1940

OTHER INFORMATION: /product= "human menin"

FEATURE:

NAME/KEY: exon

LOCATION: 1..87

FEATURE:

NAME/KEY: exon

LOCATION: 88..555

FEATURE:

NAME/KEY: exon

LOCATION: 556..764

FEATURE:

NAME/KEY: exon

LOCATION: 765..893

FEATURE:

NAME/KEY: exon

LOCATION: 894..934

FEATURE:

NAME/KEY: exon

LOCATION: 935..1022

FEATURE:

NAME/KEY: exon

LOCATION: 1023..1159

FEATURE:

NAME/KEY: exon

LOCATION: 1160..1295

```
;
; NAME/KEY: exon
; LOCATION: 1296..1460
; FEATURE:
;
; NAME/KEY: exon
; LOCATION: 1461..2764
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-380-337-1

Query Match      14.3%; Score 1312.2; DB 5; Length 2772;
Best Local Similarity 96.5%; Pred. No. 1.1e-19;
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7499 GGATGGCCAGAGCAGGGTCTGTGAGTTCCAGCCACTGGCCGGCAACCTTCTCTCACCTT 7558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1382 GGAGGGCAGTCCCACGCCCTGTGTGCACGTGGCTGGCCACCTTTCTGTGTCAGTCCCT 1441

QY 7559 GCTCTCCCACTGGCCAGGTGGCGGAGAGGTGCGCATAGTGAAGCCGAGAGCCGAGGC 7618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1442 AGCCCGTTTTGAGGGACAGGTGGCGGAGAGGTGCGCATAGTGAAGCCGAGAGCCGAGGC 1501

QY 7619 GCGGAGGCGGAGGAGCCGTGGGGCGAGGAGAGCCCGGGAAGCCGGCGGGGCCACG 7678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1502 GCGGAGGCGGAGGAGCCGTGGGGCGAGGAGAGCCCGGGAAGCCGGCGGGGCCACG 1561

QY 7679 GCGGGAGTCCCAAGCCAGAGAGCCCGCCGCCCAAGAGCCAGCACTGGACAGGGCCCT 7738
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1562 GCGGGAGTCCCAAGCCAGAGAGCCCGCCGCCCAAGAGCCAGCACTGGACAGGGCCCT 1621

QY 7739 GGCACCGGCCAGGGTGCAAGTGTGAGAGCCCGCCGCCCAAGAGCCCTCTCTGGACTGG 7798
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1622 GGCACCGGCCAGGGTGCAAGTGTGAGAGCCCGCCGCCCAAGAGCCCTCTCTGGACTGG 1681

QY 7799 CACAGCCCGAGGCCCTGAGAGTGGGAGCAGCAGCGCTCAGGTGCGCAGCACCCGACATACC 7858
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1682 CACAGCCCGAGGCCCTGAGAGTGGGAGCAGCAGCGCTCAGGTGCGCAGCACCCGACATACC 1741

QY 7859 ACCGCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGATGAAGGATGAAGGAGCT 7918
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1742 ACCGCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGATGAAGGATGAAGGAGCT 1801

QY 7919 GCTGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCA 7978
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1802 GCTGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGGCACAGTCGCA 1861

QY 7979 AGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTTCTCAA 8038
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1862 AGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTTCTCAA 1921

QY 8039 GCGGACGCGCAAGGCTCTGAACTACTGGGACTTTCGGACCCCTTGTGGGACCCAGGC 8098
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1922 GCGGACGCGCAAGGCTCTGAACTACTGGGACTTTCGGACCCCTTGTGGGACCCAGGC 1981

QY 8099 TCGGCTTGTATCCCGCACTCTGAGCGCCATGTGTCGCCCGCCAGCCCAAGGGGACAGGC 8158
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1982 TCGGCTTGTATCCCGCACTCTGAGCGCCATGTGTCGCCCGCCAGCCCAAGGGGACAGGC 2041

QY 8159 TCACCTCTACCAACCCCTAGTGTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGAT 8218
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2042 TCACCTCTACCAACCCCTAGTGTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGAT 2101

QY 8219 TTTCTCCAGTCAAGACCCAGGCTCTGCCCCAGTCTGTTAGAATATAGTCTCTCTCCC 8278
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2102 TTTCTCCAGTCAAGACCCAGGCTCTGCCCCAGTCTGTTAGAATATAGTCTCTCTCCC 2161

QY 8279 AGAATCCCGAGCGGCAATGGAAACCTCAGCTGGGTGCTTAATACCAAGCTTTTAAAGGC 8338
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2162 AGAATCCCGAGCGGCAATGGAAACCTCAGCTGGGTGCTTAATACCAAGCTTTTAAAGGC 2221

QY 8339 CCAGGCCCCAGAAACCAAGCTCTCTCGGAACCCGCTACCTTAGAGCCAGACCAAGCTT 8398
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2222 CCAGGCCCCAGAAACCAAGCTCTCTCGGAACCCGCTACCTTAGAGCCAGACCAAGCTT 2281
```

```
QY 8399 ACTCAGGGCTCCTCCAGCTTGTAGGAGCTTGTAGGCTTACACCTTAACCCAAGGAGCACA 8458
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2282 ACTCAGGGCTCCTCCAGCTTGTAGGAGCTTGTAGGCTTACACCTTAACCCAAGGAGCACA 2341

QY 8459 GTTCCCACTCCAGCCGCGGGAGCCCTAGGACCACTCAGCCCCCTAGGAGTATATTTCCGCA 8518
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2342 GTTCCCACTCCAGCCGCGGGAGCCCTAGGACCACTCAGCCCCCTAGGAGTATATTTCCGCA 2401

QY 8519 CTTCAGAATTCCATATCTTCGGAATCCAAGCTCCCTGCCCAATAAATTAATTCAGTCTCGCT 8578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2402 CTTCAGAATTCCATATCTTCGGAATCCAAGCTCCCTGCCCAATAAATTAATTCAGTCTCGCT 2461

QY 8579 TTCAGAATTTGGAATCTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8638
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2462 TTCAGAATTTGGAATCTCCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2521

QY 8639 CCGCCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8698
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2522 CCGCCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 2581

QY 8699 TCAGAGCAGGACCTCCTTCCGACCCTCTGGGAACCTCCCGAGAGTCCAGCCCATCTCGG 8758
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2582 TCAGAGCAGGACCTCCTTCCGACCCTCTGGGAACCTCCCGAGAGTCCAGCCCATCTCGG 2641

QY 8759 AGCATCCCGGAGGAAATCTCGAGAGGGTTAGGAGTGGGTGACAAGAGCTGTATCTCTTC 8818
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2642 AGCATCCCGGAGGAAATCTCGAGAGGGTTAGGAGTGGGTGACAAGAGCTGTATCTCTTC 2701

QY 8819 CTCTTTTGTACATAGATTTATTTTTCAGTTTCCAGAAAGATGAATACATTTTGTGTAATA 8878
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 2702 CTCTTTTGTACATAGATTTATTTTTCAGTTTCCAGAAAGATGAATACATTTTGTGTAATA 2761

QY 8879 AATATATAA 8887
    ||| ||| |||
Db 2762 AAAAAAAAA 2770
    ||| ||| |||

RESULT 10
US-09-948-941-125
; Sequence 125, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE OF INVENTION: WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL000798
; CURRENT APPLICATION NUMBER: US/09/948,941
; PRIOR FILING DATE: 2001-09-10
; PRIOR FILING DATE:
; SOFTWARE: FastSeq for Windows Version 4.0
; NUMBER OF SEQ ID NOS: 12618
; SEQ ID NO 125
; LENGTH: 2767
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-125

Query Match      13.7%; Score 1256.4; DB 5; Length 2767;
Best Local Similarity 96.3%; Pred. No. 1.4e-18;
Matches 1339; Conservative 0; Mismatches 46; Indels 5; Gaps 5;

QY 7499 GGATGGCCAGAGCAGGGTCTGTGAGTTCCAGCCACTGGCCGGCAACCTTCTCTCACCTT 7558
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1382 ggaaggcagtcacccacccctgtgtgcacgtgggctggggccaccccttctgtgcagtcct 1441

QY 7559 GCTCTCCCCACTGGCCAGGTCGGGAGAGGTGCGCATAGTGAAGCCGAGAGCCGAGGC 7618
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1442 aggccttttgaggagacaggtgcggcagaaggtgcgcatagtgagccgagagccgagcc 1501

QY 7619 GCGGAGGCGGAGGAGCCGTCTCTCGGAACCCGCTACCTTAGAGCCGCGGCGGCCACG 7678
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```


Db	1242		agtgacagatgaagagcagagaagtgccaccctctagtgactcaactctgtctttctctcaa	1301
Qy	8039		GGCGCAGCCAAAGCCCTCTGAACCTACTTGGGGACTTTCGGACCGCTGTGGGGACCCAGGC	8098
Db	1302		gcgagcgcmaaaagccctctgaactacttgggacctcgagccgcttgtgggaccagcgc	1361
Qy	8099		TCCG-CTTTAGTCTCCCAACTCTGAGCCCATCTGTCGCCCCAGCCCAAGGGGACAGGC	8157
Db	1362		tccgcctctagtcccccaactctgagccatgtctgtccccagcccaaggagcagcgc	1421
Qy	8158		CTCACCTCTACCCAAACCTTAGTTTCCCGGTCCCGAGTACAGTCTGTATCAAAACCACAGA	8217
Db	1422		ctcacctctaccccaaaccttaagttcccggtcccaggtacagtcgtgtatacaaccacaga	1481
Qy	8218		TTTTTCTCCAGCTCAGAACCCAGGGCTCTGCCCCAGTCGTGTAGAATATAGTCTCTTCTCC	8273
Db	1482		ttttctccagctcagaaaccagggtctgtgcccaagtctgtagaataagtgctctctctcc	1541
Qy	8278		CAGAACTCCAGCGCGCCCAATGGAACCTTCAGCTGGGTCTTAATACCACTTTTAAAGG	8337
Db	1542		cagaa tcccagcgcgccaatggaaacctca cgc tgggtccctaatca cca gtcctttaaagg	1601
Qy	8338		CCAGGCCCTAGAAACCCAAAGCTCTCTCTCGGAACCGCTCACTCTAGAGCCAGACCAACGT	8397
Db	1602		cccagccctagaacccaagctctctctcgaaaccgtcaactagagccagaccaactg	1661
Qy	8398		TACTCAGGCTCTCTCCAGCTTGTGTAGAGCTGTAGGTTTTCACCCCTTAACCAAGGGAGCAC	8457
Db	1662		tactcaggctctctccagcttgttaggagctgaggtttca ccc tta ccc caa -ggagcac	1720
Qy	8458		AGTCTCCACCTCCAGCCGGGGAGCCTAGGACCCTCAGCCCTCAGCGCCCTAGGAGTATTTCCCG	8517
Db	1721		aggtccccactccagccc -ggagcctaggaacctcagcccccttaggagtatattccgc	1779
Qy	8518		ACTTCAGAAATCCATATCTTGGCAATCCAAGCTCCCTGCCCCAATAACTTCAGTCTCTGC	8577
Db	1780		acttcagaattccatatcttgcgaatccaagctccctgcgcccaataaactcagtcctgc	1839
Qy	8578		TTCCAGAAATTTGGAATCGTAGTTTCTCTCTTCGTATCCGAGTCTGGGACACAAAC	8637
Db	1840		-tccagaatttggaaatcctagtttctctctctcgatcccgagctgtggacacaaaac	1898
Qy	8638		TCCGCCCCAGGCTATGACATCTTAGCCCGCCCTCTCTCTCAGCAAACTGSCCCCGG	8697
Db	1899		tccgccccagcctatgagcatccttgagccccgcctctctctgacgaactggccccgg	1958
Qy	8698		ATCAGACGAGACCTCCCTCTCCAGCCCTCTGGAACTCCCGAGGTCACAGGCCATCTCT	8757
Db	1959		atcagagcagacctctctccagccctctggaaacctcccagaggtccagccatctcg	2018
Qy	8758		GAGCATCCCGGAGGAATCTGCAGAGGGTTTAGAGTGGGTGACAAAGAGCTGATCTCTT	8817
Db	2019		gagcatcccggaggaaatctgcaga -ggagttaggagtggtgacaa gacctgatctctt	2077
Qy	8818		CTGTCTTTGTACATAGATTATTTTTCAGTTTCCAAAGAAAGATGAATACATTTCTTAAAA	8877
Db	2078		ctgtgtttgatcatagatttatttttcagttccaaagaagaatgaatacattttgttaaaa	2137
Qy	8978		AAAAATATAA 8887	
Db	2138		aaaaaaaaa 2147	

RESULT 12

US-60-278-561-1743

US-00-278-301-1743
; Sequence 1743, Application US/60278561

[illegible]

APPLICANT: Morris, Macdonald

```

; APPLICANT: MORRIS, MACDONALD
: APPLICANT: Lal, Preeti

```

APPLICANT: Lai, Preethi

; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the identification of sequence polymorphisms

TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms

; TITLE OF INVENTION: Polymorphisms Identified Thereby

FILE REFERENCE: GX-0012-1 P

FILE REFERENCE: SA 0012 1
; CURRENT APPLICATION NUMBER: US/60/278,561

; CURRENT AFFILIATION NUMBER: 05/0
 ; CURRENT FILING DATE: 2001-03-23

; CURRENT FILING DATE: 2001
; NUMBER OF SEQ ID NOS: 15598

```

; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PERL Program

```

: SOFTWARE: PER:
: SEO ID NO 1743

```

; SEQ ID NO 1743
; LENGTH: 3534

```

```

; LENGTH: 3534
; TYPE: DNA

```

```

; TYPE: DNA
: ORGANTSM: Homo sapiens

```

ORGANISM
FEATURE:

```
; FEATURE:
: NAME/KEY: misc feature
```

```

; NAME/KEY: misc_feature
;
; OTHER INFORMATION: Incvte ID No: 074428 3

```

; OTHER INFORMATION
US-60-278-561-1743

Query Match	13.4%	Score 1231.4	DB 8	Length 3534	
Best Local Similarity	96.2%	Pred. No. 3.4e-18			
Matches 1325	Conservative	0	Mismatches 46	Indels	Gaps 6
QY 7499	GGATGGCCAGAGCAGGCTCTGGAGTTCCAGCCACTGGCGCGCAACCTGTGCTCTCACCTT	7558			
DB					
DB 2162	ggagggcagtcaccagcctgtgtgcaagtgggctgggccaacctttcttggagatccct	2221			
QY 7559	GCTCTCCCCACTGGCCACGAGTGGCGCAGAAAGTGGCATAGTGACCGCAGAGCCGAGGC	7618			
DB					
DB 2222	aggcgcttttgaggagcaggtgcggcagaagtgcgcatagtgagccgagggccgaggc	2281			
QY 7619	GGCGAGCGCCAGAGCCGTGGGGCGAGGAAGCCCGGAAGCCGGCGCGGGCGGCCACG	7678			
DB					
DB 2282	ggccagggccgagggagccgtggggaggaagccggagggagccggcgggggccacg	2341			
QY 7679	GCGGAGTCCAAAGCCAGAGAGCCCGCGCGCCCAAGAACGACACTGGACAAGGGCCT	7738			
DB					
DB 2342	gcggagttcaagccagagagcccgccgcgccaagaagcagcactggacaaggcct	2401			
QY 7739	GGGACCGCGCCAGGGTGCAGTGTCCAGA-CCGCCCGGAGACCTCTCTGGGACTGTCCGTG	7797			
DB					
DB 2402	gggcacggccaggggtcagtgtagcagcccccccggaagcctcctgggactgcgtg	2461			
QY 7798	GCACAGCCGAGGGCCCTGAAGTGGCAGACCGGCTCAGGTCCAGCCAGCACCCCGACATCAC	7857			
DB					
DB 2462	gcacagcccgagggccctggaagtgaggcagcgctcaaggtgccaagcaccgcagcatcac	2521			
QY 7858	CACCGCGGAGGGTCCAGTGCTCACTTTCCAGGTGAGAAGATGAAGGCGATGAAGGAGC	7917			
DB					
DB 2522	cacgcgcgaggggtccagtgctcacttccagagtgagaagatgaaggcgcatgaaggagc	2581			
QY 7918	TGCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTCCAATCACGGCACAGTCGC	7977			
DB					
DB 2582	tgcctgtggccaccagatcaactctgagcgccatcaagctgcaactcacggcagtcgc	2641			
QY 7978	AAGTGCAGATGAAGAAGCAAAAGTGCACCCCTAGTGACTACACTGTGCTTTCTCTCA	8037			
DB					
DB 2642	aagtgcagatgaagaagcagaaggtgccaccctagtactacactctgtcttctctca	2701			
QY 8038	AGCGGCAGCGCAAGGCGCTCTGAACACTCTGGGACTTCGGACCGCTGTGGGGACCCAGG	8097			
DB					
DB 2702	agcggcagcgcaagggcctctgaactactgggagcttcggaccgcttggggggcccaag	2761			
QY 8098	CTCGG-CCTTGTAGTCCCCCACTCTGAGCCCATGTTCTGCCCCAGCCCAAGGGGACAGG	8156			
DB					
DB 2762	ctccgccttagtcccccaactctgagcccatgttctgtcccccaagcccaagggagacag	2821			
QY 8157	CCTCACCTCTACCCAAACCCTTAGTTCCTCCGCTCCGAGTACAGCTGTGTATCAAAACCCACG	8216			
DB					
DB 2822	cctcacctctacccaaaccttagttcccggtcccgagtcacagtcgtgtatcaaacccacg	2881			
QY 8217	ATTTTCTCCAGCTCAGAACCCAGGGCTCTGCCAGTCGTTAGATAATAGTCTCTCTCTC	8276			
DB					
DB 2882	atttctccagctcagaaccagggcctgtgccagtcgttgaataatagatctctcttc	2941			
QY 8277	CCAGAATCCAGCGCGCAATGGAACCTCACGCTGGGTCCCTAAATTACCAGCTTTTAAAG	8336			

Db	2942	 ccagaaatcccaggccggccaattggaacctcaccgcgctgggttcctaatactccagtccttaag	3001
Qy	8337	GCCCAGCCCTAGAAACCAACGACTCCTCTCGGAACCCTCACCTAGAGCCAGACCAACG	8396
Db	3002	gccagcccttagaaaaccaaactcctcctcggaaccgctcacctagagccagaccaacg	3061
Qy	8397	TTACTCAGGGCTCTCCAGCTGTGTAGAGCTGAGGTTTCACCTTAACCCAAGGGAGCA	8456
Db	3062	ttactcagggtcctctcccaagtgttaggagctggaggttccaccttaacccaa- gggagca	3120
Qy	8457	CAGGTCCCACCTCCAGCCGGGAGCCTTAGGACCACTCAGGCCCTTAGGAGTATATTCCG	8516
Db	3121	cagggtcccacctccagccc- gggagccttagaaccaactcagccccctaggagtatattccg	3179
Qy	8517	CAGTTCAGAATTCCCATATCTTCGGAATCAAAGCTCCCCCTGCACCAATAACTTCAGTCTGT	8576
Db	3180	cacttcagaattcccatatcttggaatccaagctccctgccccaaaaaacttcagtcctg	3239
Qy	8577	CTTCCAGAATTGGGAATCCTAGTTTCTCTCTCTTGATCCCGAGTCTGGGACACAAAA	8636
Db	3240	c- tccagaatttggaaatcctagtttctctctctctcgatatccccagttctgggacacaaa	3298
Qy	8637	CTCCGCCCCAGCCTATGACATCTGTAGCCCGCCCTCTTCTCTGACAAACTGGCCCGC	8696
Db	3299	ctcgcgcccgacctatgagcactcagcccgccctctcttcgacgaaactggccccg	3358
Qy	8697	GATCAGAGCAGGACCTCCCTTCGGACCTCTGTGGGAACCTCCAGAGGTTCAGCCCATCTC	8756
Db	3359	gatcagagcaggacctcccttcgcacctctgggaacctccagagg'tccagcgccatctc	3418
Qy	8757	GGAGCATCCGAGGAGAAATCTCAGAGGGGTTT'PAGAGTGGGTGAC'AAGAGCCTGATCTCT	8816
Db	3419	ggagcatcccgagggaaa'ctctcaga - gggttaggag'tgggtgacaagagcctgatctct	3477
Qy	8817	TCCTGTTTTGTACATAGATTTATTTTTCAGTTTCCAGAAAAGATGAATACATTTTGT	8873
Db	3478	tctgttttgtacatagatttatatttcagttccagaaadaataatacatacttttgt	3534

```

RESULT 13
US-60-278-561-1743/c
; Sequence 1743, Application US/60278561
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lal, Preeti
; APPLICANT: Deep, Dich
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms Using
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0012-1 P
; CURRENT APPLICATION NUMBER: US/60/278,561
; CURRENT FILING DATE: 2001-03-23
; NUMBER OF SEQ ID NOS: 15598
; SOFTWARE: PERL Program
; SEQ ID NO 1743
; LENGTH: 3534
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 074428.3
US-60-278-561-1743

```

Query Match	7.6%	Score 698.6;	DB 8;	Length 3534;
Best Local Similarity	95.5%;	Pred. No. 6.4e-08;		
Matches 719; Conservative	0;	Mismatches 34;	Indels 0;	Gaps 0;
Qy 2229	CCTGCCGGACCTCCCTCCTCCCGGCTTGAGGCGCGCGCCACCGCGCCGC	2288		
Dd 753	CCACCCTTGCCTCCCGCAGAGCTTGAATATAGTGGCGGGCGCGCGCGCTCGGTGGCTGA	694		

Qy	2289	TGGGGCTGAAGGCCGCCCAAGAGACGCTGTTCCCGTGGCGTCCATCCAGACAGCTGGTGC	2348
Db	693	AGGGGCTGAAGAGGCCGCCCAAGAGACGCTGTTCCCGCTGGCTTCCATCCAGACAGCTGGTGC	634
Qy	2349	GCCTGTTTGCTGCCGAGCTGGGCCGAGAGAGCGGAGCACTGGTGGCTCCCTTTCCTTGGTGC	2408
Db	633	GCCTGTTTGCTGCCGAGCTGGGCCGAGAGAGCGGAGCACTGGTGGCTCCCTTTCCTTGGTGC	574
Qy	2409	TGGGCTTCTGGTGGAGCATTTTCTTGGCTGTCAACCGGGTCAATCCCTTACCAACGTTCCCGAGC	2468
Db	573	TGGGCTTCTGGTGGAGCATTTTCTTGGCTGTCAACCGGGTCAATCCCTTACCAACGTTCCCGAGC	514
Qy	2469	TCACCTTCCAGCCAGCCGCCGCCCGACCCGCGCTGGCGGGCTCACTTCTCCCGTGG	2528
Db	513	TCACCTTCCAGCCAGCCGCCGCCCGACCCGCGCTGGCGGGCTCACTTCTCCCGTGG	454
Qy	2529	CCGACCTGTCTATCATCGCGCGCCCTCTATGCGCGCTTCAACGCCCCAGATCGGAGGCGCGC	2588
Db	453	CCGACCTGTCTATCATCGCGCGCCCTCTATGCGCGCTTCAACGCCCCAGATCGGAGGCGCGC	394
Qy	2589	TCGACCTGTCCCTCTATCTCTCGAAGAGGGGTGTTCTCAGCCGTCAGCTGTTGTAAGAANG	2648
Db	393	TCGACCTGTCCCTCTATCTCTCGAAGAGGGGTGTTCTCAGCCGTCAGCTGTTGTAAGAANG	334
Qy	2649	TCTCCGATGTCAATATGGAACAGCGCTCAGCGCGTCTCTACTTCAAGGATCGGGCCACATCC	2708
Db	333	TCTCCGATGTCAATATGGAACAGCGCTCAGCGCGTCTCTACTTCAAGGATCGGGCCACATCC	274
Qy	2709	AGTCCCTCTTTCAGCTTCATCAGAGTTGGAGCCCACTAGTGGGGAATCTTATCATGACC	2768
Db	273	AGTCCCTCTTTCAGCTTCATCAGAGTTGGAGCCCACTAGTGGGGAATCTTATCCATGACC	214
Qy	2769	CACCTTCTTAAACCTCCATGGTTTACAGAACCCTTTTAAAGACTGTAAAGCCTTGTGAG	2828
Db	213	CACCTTCTTAAACCTCCATGGTTTACAGAACCCTTTTAAAGACTGTAAAGCCTTGTGAG	154
Qy	2829	GTTCCGGCAGGTGTATTTTCTCTTTTTCAGTTTGGGAACCTGAAGCCCAAGAGAGGGGAAT	2888
Db	153	GTTCCGGCAGGTGTATTTTCTCTTTTTCAGTTTGGGAACCTGAAGCCCAAGAGAGGGGAAT	94
Qy	2889	GATATGCCAAAGTCACACACGGGATGGCAGGCTTGAAGTGAAGCTCGATCACTTGGCTC	2948
Db	93	GATATGCCAAAGTCACACACGGGATGGCAGGCTTGAAGTGAAGCTCGATCACTTGGCTC	34
Qy	2949	CAATCATCAACCTCACTCTGCCCCCTCAGCA	2981
Db	33	CAATCATCAACCTCACTCTGCCCCCTCAGCA	1

```

RESULT 14
US-09-864-761-22111/c
: Sequence 22111, Application US/09864761
: GENERAL INFORMATION:
: APPLICANT: Penn, Sharon G.
: APPLICANT: Rank, David R.
: APPLICANT: Hanzel, David K.
: APPLICANT: Chen, Wensheng
: TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLES
: TITLE OF INVENTION: GENE EXPRESSION ANALYSIS
: FILE REFERENCE: Aeo mica-X-1
: CURRENT APPLICATION NUMBER: US/09/864,761
: CURRENT FILING DATE: 2001-05-23
: PRIOR APPLICATION NUMBER: US 60/180,312
: PRIOR FILING DATE: 2000-02-04
: PRIOR APPLICATION NUMBER: US 60/207,456
: PRIOR FILING DATE: 2000-05-26
: PRIOR APPLICATION NUMBER: US 09/632,366
: PRIOR FILING DATE: 2000-08-03
: PRIOR APPLICATION NUMBER: GB 24263,6
: PRIOR FILING DATE: 2000-10-04
: PRIOR APPLICATION NUMBER: US 60/236,359
: PRIOR FILING DATE: 2000-09-27
: PRIOR APPLICATION NUMBER: PCT/US01/00666

```

APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY

;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00667
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00664
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00669
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00665
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00668
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00663
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00662
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00661
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: PCT/US01/00670
;; PRIOR FILING DATE: 2001-01-30
;; PRIOR APPLICATION NUMBER: US 60/234,687
;; PRIOR FILING DATE: 2000-09-21
;; PRIOR APPLICATION NUMBER: US 09/608,408
;; PRIOR FILING DATE: 2000-06-30
;; PRIOR APPLICATION NUMBER: US 09/774,203
;; PRIOR FILING DATE: 2001-01-29
;; NUMBER OF SEQ ID NOS: 49117
;; SOFTWARE: Anomax Sequence Listing Engine vers. 1.1
;; SEQ ID NO 22111
;; LENGTH: 639
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; OTHER INFORMATION: MAP TO AC000134.14
;; OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.7
;; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
;; OTHER INFORMATION: EXPRESSED IN LUNG, SIGNAL = 1.2
;; OTHER INFORMATION: EXPRESSED IN HEART, SIGNAL = 1.8
;; OTHER INFORMATION: EXPRESSED IN BONE MARROW, SIGNAL = 2
;; OTHER INFORMATION: EXPRESSED IN ADULT LIVER, SIGNAL = 1.5
;; OTHER INFORMATION: EXPRESSED IN HELA, SIGNAL = 2.1
;; OTHER INFORMATION: EXPRESSED IN FETAL LIVER, SIGNAL = 2.5
;; OTHER INFORMATION: EST_HUMAN HIT: BF526328.1, EVALU0 0.00e+00
;; OTHER INFORMATION: NT HIT: 093237.1, EVALU0 0.00e+00
;; OTHER INFORMATION: SWISSPROT HIT: 000255, EVALU0 2.00e-43
US-09-864-761-22111

Query Match 7.0%; Score 639; DB 7; Length 639;
Best Local Similarity 100.0%; Pred. No. 3.4e-06;
Matches 639; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 7422 CAGACTTTTCTGGCCAGGGGAGCATCTCCCATCCCTTCGTCGGTCCGATGGAGTGGAG 7481
Db 639 CAGACTTTTCTGGCCAGGGGAGCATCTCCCATCCCTTCGTCGGTCCGATGGAGTGGAG 580
QY 7482 ACCCCCTGGGTGGATGGATGCCAGAGCAGGGTCTCTGGAGTTCAGGCACACTGGCCGGC 7541
Db 579 ACCCCCTGGGTGGATGGATGCCAGAGCAGGGTCTCTGGAGTTCAGGCACACTGGCCGGC 520
QY 7542 AACCTTGTCTACCTTGTCTCTCCCACTGGCCAGGTGCGGCAGAGGTGCGCATAGTG 7601
Db 519 AACCTTGTCTACCTTGTCTCTCCCACTGGCCAGGTGCGGCAGAGGTGCGCATAGTG 460
QY 7602 AGCCGAGAGCCGAGCCGAGCCGAGGAGCCGTGGGGGAGAGCCGCGGGAAGCC 7661
Db 459 AGCCGAGAGCCGAGCCGAGCCGAGGAGCCGTGGGGGAGAGCCGCGGGAAGCC 400
QY 7662 CGCGCGGGGGCCACCGCGGGAGTCCAAAGCAGAGAGCCCGCCGCGCCCAAGAGCCA 7721
Db 399 CGCGCGGGGGCCACCGCGGGAGTCCAAAGCAGAGAGCCCGCCGCGCCCAAGAGCCA 340
QY 7722 GCACTGGACAAGGGCTGGGGCAGCCGCGGAGGTGCAGTGTCAAGACCCCGCCGGAAGCCT 7781

Db 339 GCACTGGACAAGGGCTGGGGCAGCCAGGGTGCAGTGTCAAGACCCCGCCGGAAGCCT 280
QY 7782 CTTGGGACTGTGCTGGCAGACAGCCCGAGCCCTGAAGTGTGGCAGCAGCGTCAAGTGGCCA 7841
Db 279 CTTGGGACTGTGCTGGCAGACAGCCCGAGCCCTGAAGTGTGGCAGCAGCGTCAAGTGGCCA 220
QY 7842 GCACCCGCGCAGCATCACACCCGCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGATG 7901
Db 219 GCACCCGCGCAGCATCACACCCGCGGAGGGTCCAGTGTCTCACTTTCCAGAGTGAAGATG 160
QY 7902 AAGGGCATGAAGAGTGTCTGGTGGCCACCAAGATCAATCGAGCGCCATCAAGCTGCA 7961
Db 159 AAGGGCATGAAGAGTGTCTGGTGGCCACCAAGATCAATCGAGCGCCATCAAGCTGCA 100
QY 7962 CTCACGGCAGCTGCAAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTAC 8021
Db 99 CTCACGGCAGCTGCAAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTAC 40
QY 8022 ACTCTGTCTTCTCTCAAGCGGCGAGCGCAAGGCTCTGA 8060
Db 39 ACTCTGTCTTCTCTCAAGCGGCGAGCGCAAGGCTCTGA 1
RESULT 15
US-09-948-941-1028
; Sequence 1028, Application US/09948941
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH CANCER, METHODS OF DETECTION AND USES THEREOF
; CURRENT APPLICATION NUMBER: US/09/948,941
; PRIOR FILING DATE: 2001-09-10
; PRIOR APPLICATION NUMBER: 60/231,328
; NUMBER OF SEQ ID NOS: 12618
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1028
; LENGTH: 601
; TYPE: DNA
; ORGANISM: Human
US-09-948-941-1028
Query Match 6.5%; Score 600.6; DB 5; Length 601;
Best Local Similarity 99.8%; Pred. No. 1.9e-05;
Matches 600; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
QY 6964 GCCTTCAGTTTTCATGGCCAGACTCCCTCCCTCAGCTGAGGGGTGGAGGTAGGGATGTAC 7023
Db 1 gccttcagtttctcatggccagactccctccctcagctgaggggtgaggttaggtggtac 60
QY 7024 GTCTCGCTATGGATGGCTTTTATAAGAAAGAGAGTTCTAAGATGTTCACCACTAT 7083
Db 61 gtctcgctatggatggcttttataagaaagagagttcttaagaatgttcccaacctat 120
QY 7084 GCTTACCTTTCTGAGAGCCAGGGGTCTTTCCTAGTGGGGGCTCGCCCTGTGCTCT 7143
Db 121 gcttacctttctgagagccaggggtctttcctagtgaggggttgcctctgacctct 180
QY 7144 GCTAAGGGGTGAGTAAGAGACTGATCTGTGCTCCCTCCCTTCCTCCAGGCGACCCA 7203
Db 181 gctaaggggtgagtaagagactgactgtgacctctccctctccctccctccctccctcc 240
QY 7204 GAGCCAAAGGTTCGCCCTCCAGGACCTTGTCTGCTGCGCCACTGCTGCGATTTACGA 7263
Db 241 gagccaaaggttcgccctccagagacctgtgtcttcgccccactgtgctgacttactaga 300
QY 7264 CGGCATCTGCAATGGAGGAGGGCAGTCCACGCTCTGCTGCACCTGGCTGGGCGCAC 7323
Db 301 yggcatctgcaaatggagagggaggtccacgctgtgctgcaagtggtggtggtggtggtggt 360
QY 7324 CTTTCTGTGTCTCCCTAGGCGGTTTTTGGAGGAGAGGTGAGGAGACAGCTGCACAGGT 7383

```
|||||
Db 361 ctttctgtgcagtccttagcgcttttgaggacaggtgaggacagctgcacagaggt 420
QY 7384 CTGGGCACTACAGGTGGTGACAGCACGGGCTTGTGACACTTTTCTGCCCCAGGGGC 7443
Db 421 ctgggcactacaggtggtgacagcagccacgggcttgcagactttctgcccaggggc 480
QY 7444 AGCATCTGCCCATCCCTTTCGGTGCCGATGGGACTGAGACCCCTGGGTGGGATGGGATG 7503
Db 481 agcatctgcccatcccttcggtgccgatggactggagacccctgggtgggatgggatg 540
QY 7504 GCCAGAGCAGGGTCTCTGAGTTCAGCCACTGCGCCGCAACCTTGCTCTCACCTTGCTCT 7563
Db 541 gccagagcagggctcctggagtccagcactggccggcaaccttgctctcaccttgctct 600
QY 7564 C 7564
Db 601 c 601
```

Search completed: October 14, 2001, 08:34:02
Job time: 17247 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 07:39:50 ; Search time 13115.4 seconds
(without alignments)
10342.904 Million cell updates/sec

Title: US-09-380-337-3

Perfect score: 9180

Sequence: 1 CTGCTCTTGAACCTCTGGCC.....AGCCACGGGGCGCCGCCCG 9180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1415048 seqs, 7388405095 residues

Total number of hits satisfying chosen parameters: 28310096

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Pending_Patents_NA_Main.*
1: /cgn2_6/ptodata/1/pna/PCTUS_COMB.seq.*
2: /cgn2_6/ptodata/1/pna/US06_COMB.seq.*
3: /cgn2_6/ptodata/1/pna/US07_COMB.seq.*
4: /cgn2_6/ptodata/1/pna/US080_COMB.seq.*
5: /cgn2_6/ptodata/1/pna/US081_COMB.seq.*
6: /cgn2_6/ptodata/1/pna/US082_COMB.seq.*
7: /cgn2_6/ptodata/1/pna/US083_COMB.seq.*
8: /cgn2_6/ptodata/1/pna/US084_COMB.seq.*
9: /cgn2_6/ptodata/1/pna/US085_COMB.seq.*
10: /cgn2_6/ptodata/1/pna/US086_COMB.seq.*
11: /cgn2_6/ptodata/1/pna/US087_COMB.seq.*
12: /cgn2_6/ptodata/1/pna/US088_COMB.seq.*
13: /cgn2_6/ptodata/1/pna/US089_COMB.seq.*
14: /cgn2_6/ptodata/1/pna/US090_COMB.seq.*
15: /cgn2_6/ptodata/1/pna/US091_COMB.seq.*
16: /cgn2_6/ptodata/1/pna/US092_COMB.seq.*
17: /cgn2_6/ptodata/1/pna/US093_COMB.seq.*
18: /cgn2_6/ptodata/1/pna/US094_COMB.seq.*
19: /cgn2_6/ptodata/1/pna/US095A_COMB.seq.*
20: /cgn2_6/ptodata/1/pna/US095B_COMB.seq.*
21: /cgn2_6/ptodata/1/pna/US095C_COMB.seq.*
22: /cgn2_6/ptodata/1/pna/US095D_COMB.seq.*
23: /cgn2_6/ptodata/1/pna/US096A_COMB.seq.*
24: /cgn2_6/ptodata/1/pna/US096B_COMB.seq.*
25: /cgn2_6/ptodata/1/pna/US096C_COMB.seq.*
26: /cgn2_6/ptodata/1/pna/US096D_COMB.seq.*
27: /cgn2_6/ptodata/1/pna/US096E_COMB.seq.*
28: /cgn2_6/ptodata/1/pna/US097A_COMB.seq.*
29: /cgn2_6/ptodata/1/pna/US097B_COMB.seq.*
30: /cgn2_6/ptodata/1/pna/US097C_COMB.seq.*
31: /cgn2_6/ptodata/1/pna/US098_COMB.seq.*
32: /cgn2_6/ptodata/1/pna/US099_COMB.seq.*
33: /cgn2_6/ptodata/1/pna/US099A_COMB.seq.*
34: /cgn2_6/ptodata/1/pna/US099B_COMB.seq.*
35: /cgn2_6/ptodata/1/pna/US099C_COMB.seq.*
36: /cgn2_6/ptodata/1/pna/US099D_COMB.seq.*
37: /cgn2_6/ptodata/1/pna/US099E_COMB.seq.*
38: /cgn2_6/ptodata/1/pna/US099F_COMB.seq.*
39: /cgn2_6/ptodata/1/pna/US099G_COMB.seq.*
40: /cgn2_6/ptodata/1/pna/US099H_COMB.seq.*
41: /cgn2_6/ptodata/1/pna/US099I_COMB.seq.*
42: /cgn2_6/ptodata/1/pna/US099J_COMB.seq.*
43: /cgn2_6/ptodata/1/pna/US099K_COMB.seq.*

44: /cgn2_6/ptodata/1/pna/US6012_COMB.seq.*
45: /cgn2_6/ptodata/1/pna/US6013_COMB.seq.*
46: /cgn2_6/ptodata/1/pna/US6014_COMB.seq.*
47: /cgn2_6/ptodata/1/pna/US6015_COMB.seq.*
48: /cgn2_6/ptodata/1/pna/US6016_COMB.seq.*
49: /cgn2_6/ptodata/1/pna/US6017_COMB.seq.*
50: /cgn2_6/ptodata/1/pna/US6018_COMB.seq.*
51: /cgn2_6/ptodata/1/pna/US6019_COMB.seq.*
52: /cgn2_6/ptodata/1/pna/US6020_COMB.seq.*
53: /cgn2_6/ptodata/1/pna/US6021_COMB.seq.*
54: /cgn2_6/ptodata/1/pna/US6022_COMB.seq.*
55: /cgn2_6/ptodata/1/pna/US6023_COMB.seq.*
56: /cgn2_6/ptodata/1/pna/US6024_COMB.seq.*
57: /cgn2_6/ptodata/1/pna/US6025_COMB.seq.*
58: /cgn2_6/ptodata/1/pna/US6026_COMB.seq.*
59: /cgn2_6/ptodata/1/pna/US6027_COMB.seq.*
60: /cgn2_6/ptodata/1/pna/US6028_COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1425.2	15.5	2770	22	US-09-577-408-136
2	1365.2	14.9	2121	22	US-09-577-408-6141
3	1239.4	13.5	2840	49	US-60-172-360-22297
4	1006.8	11.0	32768	53	US-60-213-178-196
5	749.4	8.2	2196	12	US-08-865-337-2
6	746.4	8.1	812	12	US-08-865-337-5
7	706.2	7.7	784	29	US-09-726-211-1570
8	639	7.0	639	1	PCT-US01-00663-18748
9	600.2	6.5	632	49	US-60-177-646-792
10	522.2	5.7	603	20	US-09-539-800-12611
11	500	5.4	513	51	US-60-196-710-1772
12	486.2	5.3	551	21	US-09-540-229-133603
13	480.2	5.2	510	49	US-60-177-646-29
14	471.4	5.1	538	12	US-08-865-337-3
15	452	4.9	464	1	PCT-US01-00663-5551
16	417.2	4.5	549	27	US-09-698-013-349
17	412.2	4.5	4227	22	US-09-577-408-7026
18	402	4.4	443	28	US-09-716-990-1013
19	393	4.3	405	17	US-09-332-782-18044
20	393	4.3	405	19	US-09-515-694-18044
21	392.8	4.3	102818	54	US-60-226-178-2393
22	392.8	4.3	102818	55	US-60-233-468-2393
23	389.8	4.2	24167	1	PCT-US01-01334-8463
24	387.8	4.2	227949	54	US-60-226-178-2425
25	387.8	4.2	227949	55	US-60-233-468-2425
26	387.6	4.2	32768	53	US-60-213-162-64
27	387.4	4.2	46629	56	US-60-245-222-40
28	386	4.2	32768	53	US-60-212-655-53
29	383.8	4.2	32768	53	US-60-212-350-4
30	383.8	4.2	32768	53	US-60-213-248-12
31	383.2	4.2	32148	1	PCT-US01-01339-6906
32	382.6	4.2	30716	55	US-60-230-435-46
33	380.6	4.1	73052	56	US-60-243-780-4
34	379	4.1	32768	53	US-60-213-178-119
35	378.6	4.1	32768	54	US-60-229-515-290
36	376.4	4.1	100448	57	US-60-254-168-48
37	375.8	4.1	20009	52	US-60-206-046-15
38	373.4	4.1	32768	54	US-60-229-514-27
39	372.2	4.1	30996	53	US-60-212-656-136
40	371.2	4.0	190532	53	US-60-212-664-240
41	370.8	4.0	83750	56	US-60-245-201-48
42	369.8	4.0	86859	56	US-60-248-798-36
43	369.6	4.0	174127	53	US-60-212-769-282
44	369.2	4.0	8018	1	PCT-US01-01324-4917
45	369.2	4.0	8018	1	PCT-US01-01329-3322

ALIGNMENTS

```
RESULT
US-09-577-408-136
; Sequence 136, Application US/09577408
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Tillinghast, John
; APPLICANT: Sinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: Novel Nucleic Acids and
; FILE REFERENCE: 792
; CURRENT APPLICATION NUMBER: US/09/577,408
; CURRENT FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 8502
; SOFTWARE: pt_gct_genes Version 1.0
; SEQ ID NO 136
; LENGTH: 2770
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (979)...(1033)
; OTHER INFORMATION: similar to g1182507 in the genepept database release 115,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-577-408-136

Query Match 15.5%; Score 1425.2; DB 22; Length 2770;
Best Local Similarity 99.18; Pred. No. 1.5e-103;
Matches 1433; Conservative 0; Mismatches 13; Indels 0; Gaps 0;

Qy 2288 ATGGGGCTGAAGCGCGCCAGAGAAGCGCTGTCCCGCTGCCTCCATCGACGACGTGGTG 2347
Db 1 atggggctgaagcgcgccagagaagcgctgtcccgctgcctccatcgacgacgtggtg 60

Qy 2348 CGCCTGTTGTGCTGGAGCTGGCGGAGGAGCGGACCTGGTGGTCTCTTTCCTTGGTG 2407
Db 61 cgcctgttgtgctggagctggcgagagagcgagcgacctgggtcctcttcttggtg 120

Qy 2408 CTGGGCTTCGTGGAGCATTTCTGTGCTGCAACCGCGTCATCCCTACCAAGCTTCCCGAG 2467
Db 121 ctgggcttcgtggagcatttctggtgctcaaccgctcatccctaccacggttcccgag 180

Qy 2468 CTCACCTTCCAGCCCGCCCGCCCGCCCGCTGCGCGCTCACCCTACTTTCCCGCTG 2527
Db 181 ctcaaccttccagcccgcccgcccgcccgcccgctgcgcgctcaacctactttcccgctg 240

Qy 2528 GCCGACCTGTCTATCATCGCGCGCTCTATGCCGCTTACCGCGCCAGATCCGAGCGCC 2587
Db 241 gccgacctgtctatcatcgcgccgctctatgcccgcttcaaccgctcagcgagcgcc 300

Qy 2588 GTCGACCTGTCCCTCTATCTCGAGAAGGGGTGTCCTCAGCCGCTGAGCTGGTGAAGAAG 2647
Db 301 gtcgacctgtccctctatcatcgcgccgctctatgcccgcttcaaccgctcagcgagcg 360

Qy 2648 GTCTCCGATGTCATATGAACAGCCTCAGCGCTCCTTACTTCAAGGATCGGGCCACATC 2707
Db 361 gtctccgatgtcataatgaacagcctcagcgctcctacttcaagatcgggccacacatc 420

Qy 2708 CAGTCCCTCTTCAGCTTTCATCAGCTTGGAGCCCGAGTGGGAATCTTATCCATGAC 2767
Db 421 cagtccctcttcagcttctcatcacaggttgagcccgagtagtgggaaatctatccatgac 480

Qy 2768 CCACCTTCTTCAAAACCCCTCCATGTTTACAGAACCCCTTTTAAAGAACTGTAAGCCCTTGA 2827
Db 481 ccacttcttcaaaaacctccatggtttacagaaccttttaagaacctgtaagccttgga 540

Qy 2828 GGTTCGGCAGGTGTTATTTTCTCTTTGCGAGTTGGGAAACTGAACCCCGAGAGGGGAAA 2887
```

```
Db 541 ggttcgaggtgttatttctctcttgcagttggaaactggaagccagagaggggaaa 600
Qy 2888 TGATATGCCAAAGTCACACACGGGATGCGAGGCTTGGAAAGTGAAGCCTGATCACATTGGCT 2947
Db 601 tgaatgcccgaagtcacacacgcatggcagggctggaagtcgaagcctgaacactlggct 660
Qy 2948 CCAATCATCAACCTCAGCTCTGCCCTCAGCACCCGCCACCCCTTGGCCACTGAACAGCTA 3007
Db 661 ccaaatcatcaacctcaacctctgccccctcagcacccttggccactgaacagcta 720
Qy 3008 CAGGAGTTCTAAGCATGAGACACAGAGGGCGGCAGCATTTAGGGGCGACAGAGATGAA 3067
Db 721 caggagttctaagcatgagacacagagggcgagcagatttaggggcaagaagatgaa 780
Qy 3068 ATTGGGCTGCATTTGAGGCAGCTTAAACAAATAATGGCTATGAGATTTTATTTT 3127
Db 781 attgggctgcatttgagggcagcttaaacaaaaataatggaatctcttcttcttctt 840
Qy 3128 TTTTGTGAGACAGGGTCTCACTCTGTCCCAGGCTGGAGTGCAGTGTGTGATCATGG 3187
Db 841 ttttttgagacaggggtctcaactctgtccccaggtgagtgagtgatgacaggg 900
Qy 3188 CTCAGTCAGCCTCAGTCTCCCTGGGCTCAGAGATCTCCAACTCAGCCCTCCTGAGTAG 3247
Db 901 ctcaactgcagcctcagctctcctctgggtcagagatctctccaaacctcagcctcctgag 960
Qy 3248 CTGAGAGTACAGGATGACGCTGGTGGTGTGCTGCTTAATTTTGTATTTTGTATGAGAT 3307
Db 961 ctgagagtacaggcagctgacctggtgctggttaatttttcttcttcttcttcttctg 1020
Qy 3308 GGTGTCTCACTATGTGGCCAGAGCTGGTCTTGAACCTCTGGGCTCAAGTGTCTGCCGCG 3367
Db 1021 ggtgtcactatgtggccagactggtctgaaactcttgggtcgaagtgtctgtcccg 1080
Qy 3368 CTCAGTCTCCAAATGCTGGGATTTACAGGTGTGAGCCACCGCAACTGTGTGGCTATGAA 3427
Db 1081 ctcaactctcccaaatgctgggattacaggtgtgagccaccgcaactggtggcctatgaaa 1140
Qy 3428 ATTTTGTGCTTTCAGACGCGCTCCTCAGTCTGCGCCAGGCTGGAGTGCAGTGGTGC 3487
Db 1141 atttttcttcttctcagacggcgctcactctgctgcccaggtcgagtgagcg 1200
Qy 3488 AATCTGGCTCACTGCAAGCTCTGCCCTCCTCTTTCATGCCATCTCTCTGCTCTGCTG 3547
Db 1201 aatctcggtcactgcaagctctgctcctctgtcttctcaatctctctgctcctgct 1260
Qy 3548 CAGCCTCTGTAGTAGTGGGACTACAGGAGCCTGCCACCATGCCCTGCTTAATTTT 3607
Db 1261 cagcctctgtagtagtgaggactacaggtgcccgcacccgcccggctaaattttt 1320
Qy 3608 GGATTTTGTAGTAGACAGAGGTTTCAACCATGTTAGCCAGGATGCTCGATCTCCTGACC 3667
Db 1321 gtatttttagtagaacaggggtttcaacctgttagcagaggtggtctcgatctcctgacc 1380
Qy 3668 TCGTGAATCCCGCCCGCTTGGCTTCCCAAAGTCTCGGGATTACAGCGCTGAGCCACCGCAC 3727
Db 1381 tctgtatccacccgctcggtcctccaaaagtgtctgggtattacagggcgtgagccaccg 1440
Qy 3728 CTGGTC 3733
Db 1441 ccggcc 1446
```

```
RESULT
US-09-577-408-6141/c
; Sequence 6141, Application US/09577408
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Tillinghast, John
; APPLICANT: Sinku, Ankura
; APPLICANT: Liu, Chenghua
; APPLICANT: Drmanac, Radoje T.
```

```

; ACCRION: (1005).....(1000)
;
; OTHER INFORMATION: similar to gi3002527 in the genepept database release 115,
; OTHER INFORMATION: Run with FASTXY 3.3t00, default parameters
US-09-577-408-6141

```

QY 330/ TGGTGTCTCACTATGATGGCCAGACTGGTCTTGAACCTCTGGGGCTCAAGTGATCTGCCCG 3360

Db 1101 TGGTGTCTCACTATGATGGCCAGACTGGTCTTGAACCTCTGGGGCTCAAGTGATCTGCCCG 1042

[illegible]

```

RESULT      3
US-60-172-360-22297
; Sequence 22297, Application US/60172360
; GENERAL INFORMATION:
; APPLICANT: Morris, MacDonald
; APPLICANT: Lai, Preeti
; APPLICANT: Diep, Dinh
; TITLE OF INVENTION: Method for the Identification of Sequence Polymorphisms
; TITLE OF INVENTION: Polynucleotide Sequence Databases, and Single Nucleotide
; FILE REFERENCE: GX-0007 P
; CURRENT APPLICATION NUMBER: US/60/172,360
; CURRENT FILING DATE: 1999-12-16
; NUMBER OF SEQ ID NOS: 29838
; SOFTWARE: PERL Program
; SEQ ID NO 22297
; LENGTH: 2840
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No: 074428.3
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 2836
; OTHER INFORMATION: a, t, c, g, or other
US-60-172-360-22297

```

Query Match	11.0%;	Score 1006.8;	DB 53;	Length 32768;
Best Local Similarity	99.2%;	Pred. No. 3e-71;		
Matches 1074;	Conservative	0;	Mismatches 3;	Indels 6; Gaps 6;
Qy 8034	CTCAAGCGCGCAGCGCAAGGCCCTCTGAAC	TACTGGGGACTTCGGACCGCTGTGTGGGGACC	8093	
Db 32768	CTCAAGCGCGCAGCGCAAGGCCCTCTGAAC	TACTGGGGACTTCGGACCGCTGTGTGGGGACC	32709	
Qy 8094	CAGGCTCGG - CCTTAGTCCCGCAACTCTGAGGCCATGTCTTGCCCGCCAGCCCCAAGGGGA	8152		
Db 32708	CAGGCTCGCGCCTTAGTCCCGCAACTCTGAGGCCATGTCTTGCCCGCCAGCCCCAAGGGGA	32649		
Qy 8153	CAGGCTCACCTCTACCCAAACCCCTAGGTCGCCGTCGCCGAGTACAGTCTGTATCAAAACC	8212		
Db 32648	CAGGCTCACCTCTACCCAAACCCCTAGGTCGCCGTCGCCGAGTACAGTCTGTATCAAAACC	32589		
Qy 8213	CACGATTTTCTCCAGCTCAGAACCCAGGGCTCTTGCCCGCAGTCGTTAGAAATATAGTCTCT	8272		
Db 32588	CACGATTTTCTCCAGCTCAGAACCCAGGGCTCTTGCCCGCAGTCGTTAGAAATATAGTCTCT	32529		

```
QY 8273 TCTCCAGAAATCCACGCGCCCAATGAAGAAACCTTCACGCTGGGTCTTAATTACAGCTTTT 8332
|||||
Db 32528 TCTCCAGAAATCCACGCGCCCAATGAAGAAACCTTCACGCTGGGTCTTAATTACAGCTTTT 32469
QY 8333 AAAGGCGCGAGCCCTAGAACCCCAAGCTCTCTCGGAACCGCTACCTAGAGCCAGACC 8392
|||||
Db 32468 AAAGGCGCGAGCCCTAGAACCCCAAGCTCTCTCGGAACCGCTACCTAGAGCCAGACC 32409
QY 8393 AACGTTACTCAGGCGCTCTCCACGCTTGTAGGAGCTGAGGTTTCAACCTTAACCCAGGG 8452
|||||
Db 32408 AAGTTACTCAGGCGCTCTCCACGCTTGTAGGAGCTGAGGTTTCAACCTTAACCCAA-GG 32350
QY 8453 AGCACAGGTCCCACTCCACGCGCGGAGCCCTAGGACCACTACGCCCCCTAGGAGTATATT 8512
|||||
Db 32349 AGCACAGGTCCCACTCCACGCGCGGAGCCCTAGGACCACTACGCCCCCTAGGAGTATATT 32291
QY 8513 TCCGCGACTTCAGAAATCCATATCTTCGGAATCCAAAGCTCCCTGCCCAATAACTTCAGT 8572
|||||
Db 32290 TCCGCGACTTCAGAAATCCATATCTTCGGAATCCAAAGCTCCCTGCCCAATAACTTCAGT 32231
QY 8573 CTGCTTCCAGAAATTTGAAATCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8632
|||||
Db 32230 CTGCTTCCAGAAATTTGAAATCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 32172
QY 8633 AAAATCCGCGCGCGAGCTATGAGCATCTGAGCCCCCGCCCTCTCTCTCTCTCTCTCTCTCT 8692
|||||
Db 32171 AAAATCCGCGCGCGAGCTATGAGCATCTGAGCCCCCGCCCTCTCTCTCTCTCTCTCTCTCT 32112
QY 8693 CCGGATCAGAGCAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8752
|||||
Db 32111 CCGGATCAGAGCAGACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 32052
QY 8753 TCTCGAGCATCCCGGAGGAAATCTCGAGAGGGTTTGGAGTGGGTGACAAAGAGCCTGAT 8812
|||||
Db 32051 TCTCGAGCATCCCGGAGGAAATCTCGAGAGGGTTTGGAGTGGGTGACAAAGAGCCTGAT 31993
QY 8813 CTCTCTCTCTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAAGATGATATATTTTGT 8872
|||||
Db 31992 CTCTCTCTCTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAAGATGATATATTTTGT 31933
QY 8873 TAAAAAATATAAGCGCAAGTCCATGTTTATCTCGGAAATTTGGGATGGGGCGGGAG 8932
|||||
Db 31932 TAAAAAATATAAGCGCAAGTCCATGTTTATCTCGGAAATTTGGGATGGGGCGGGAG 31873
QY 8933 TGGAGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8992
|||||
Db 31872 TGGAGCGCCCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 31813
QY 8993 GAGCGGAGCAGAGTGGCGGGGGAAGGACGTAGGCTCCGCGCCCGCCCTTGGGGCTTCCC 9052
|||||
Db 31812 GAGCGGAGCAGAGTGGCGGGGGAAGGACGTAGGCTCCGCGCCCGCCCTTGGGGCTTCCC-CC 31754
QY 9053 CCGCGCGCGAGGGCGGCTCCCGCGGGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 9112
|||||
Db 31753 CCGCGCGCGAGGGCGGCTCCCGCGGGGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 31694
QY 9113 NGG 9115
Db 31693 CGG 31691
```

RESULT 5

```
US-08-865-337-2
; Sequence 2, Application US/08865337
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASTIA TYPE 1
; TITLE OF INVENTION: PROTEIN
; NUMBER OF SEQUENCES: 6
```

```
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2196 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT14
; CLONE: Consensus
US-08-865-337-2
```

Query Match 8.2%; Score 749.4; DB 12; Length 2196;
Best Local Similarity 97.1%; Pred. No. 1.e-50;
Matches 775; Conservative 4; Mismatches 17; Indels 2; Gaps 2;

```
QY 7683 GAGTCAAGCCAGGAGGCGCGCCGCGCCCAAGAACGACGACTGGACAGGCGCTGGGC 7742
|||||
Db 1381 GAGNCAAGCCAGGAGGCGCGCCGCGCCCAAGAACGACGACTGGACAGGCGCTGGGC 1440
QY 7743 ACCGGCCAGGCGTCAAGGTGTCAGGACCCCGCGGAGGCGCTCTCTGGGACTGTCTGGCACA 7802
|||||
Db 1441 ACCGGCCAGGCGCGAGTGTCTAGGACCCCGCGGAGGCGCTCTCTGGGACTGTCTGGCACA 1500
QY 7803 GCCCGAGGCGCTCAAGGTGTCAGGACGCGGCTCAGGTGCCAGCAGCCCGCAGCATCACACCG 7862
|||||
Db 1501 GCCCGAGGCGCTCAAGGTGTCAGGACGCGGCTCAGGTGCCAGCAGCCCGCAGCATCACACCG 1560
QY 7863 CCGGAGGCTCCAGTGTCTACTTTCCAGAGTGAGAAGATGAAGGCGCATGAAGGAGCTGCTG 7922
|||||
Db 1561 CCGGAGGCTCCAGTGTCTACTTTCCAGAGTGAGAAGATGAAGGCGCATGAAGGAGCTGCTG 1620
QY 7923 GTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGCGCACAGTCCGAAGTG 7982
|||||
Db 1621 GTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCACGCGCACAGTCCGAAGTG 1680
QY 7983 CAGATGAAGAGCAGAAAGTGTCCACCCCTAGTACTACTACTCTGTCTTTCTCTCAAGCGG 8042
|||||
Db 1681 CAGATGAAGAGCAGAAAGTGTCCACCCCTAGTACTACTACTCTGTCTTTCTCTCAAGCGG 1740
QY 8043 CAGCGAAAGCGCTCTGAAGTCTCTGGGAGTCTCGGACCGCTTGTGGGACCCAGGCTCCG 8102
|||||
Db 1741 CAGCGAAAGCGCTCTGAAGTCTCTGGGAGTCTCGGACCGCTTGTGGGACCCAGGCTCCG 1800
QY 8103 -CCTTAGTCCCCCAACTCTGTAGCCCCATGTTCTCCCCCAGCCCCAGGGGACAGGCGCTCA 8161
|||||
```

Db 1801 CCCTAGTCCCACTCTGAGCCCTGTTCTGCCCCCAGCCCAAGGGGAGAGGCTCA 1860
Qy 8162 CTTCTACCAACCTAGTTTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGGATTTT 8221
Db 1861 CTTCTACCAACCTAGTTTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGGATTTT 1920
Qy 8222 CTTCTACCAACCTAGTTTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGGATTTT 8281
Db 1921 CTTCTACCAACCTAGTTTCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGGATTTT 1980
Qy 8282 ATCCAGCCGCGCAATGAAACCTCAGCTGGTCCCTAATTTACAGTCTTTAAAGGCCCA 8341
Db 1981 ATCCAGCCGCGCAATGAAACCTCAGCTGGTCCCTAATTTACAGTCTTTAAAGGCCCA 2040
Qy 8342 GCCCTAGAAACCAAGCTCTCTCGG-AACCGCTCACCTAGAGCCAGACCAACGTTAC 8400
Db 2041 GCCCTAGAAACCAAGCTCTCTCGG-AACCGCTCACCTAGAGCCAGACCAACGTTAY 2100
Qy 8401 TCAGGGTCTCCAGCTGTAGAGCTGAGGTTTCACCTTAACCCAAAGGAGCACAGG 8460
Db 2101 TCAGGGTCTCCAGCTGTAGAGCTGAGGTTTCACCTTAACCCAAAGGAGAACNGG 2160
Qy 8461 TCCACCTCCAGCCCGG 8478
Db 2161 TCCACCTCCAGCCCGG 2178

RESULT 6

US-08-865-337-5

; Sequence 5, Application US/08865337

; GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice

; APPLICANT: Covitz, Peter

; APPLICANT: Tang, Y. Tom

; APPLICANT: Murry, Lynn E.

; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE 1

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Incyte Pharmaceuticals, Inc.

; STREET: 3174 Porter Drive

; CITY: Palo Alto

; STATE: CA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette

; COMPUTER: IBM Compatible

; OPERATING SYSTEM: DOS

; SOFTWARE: FastSeq for Windows Version 2.0

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/865,337

; FILING DATE: Herewith

; CLASSIFICATION: 530

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; ATTORNEY/AGENT INFORMATION:

; NAME: Billings, Lucy J.

; REGISTRATION NUMBER: 36,749

; REFERENCE/DOCKET NUMBER: PF-0305 US

; TELEPHONE: 415-855-0555

; TELEFAX: 415-845-4166

; TELEX:

; INFORMATION FOR SEQ ID NO: 5:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 812 base pairs

; TYPE: nucleic acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; IMMEDIATE SOURCE:

; LIBRARY: BLADTUT02

; CLONE: 1313372

US-08-865-337-5

Query Match 8.1%; Score 746.4; DB 12; Length 812;

Best Local Similarity 97.2%; Pred. No. 2.7e-50;

Matches 772; Conservative 4; Mismatches 16; Indels 2; Gaps 2;

Qy 7687 CCAAGCCAGAGGAGCCCGCCGCGCCAGAGCCAGCACTGGACAAGGGCTTGGGCACCG 7746

Db 1 CCAAGCCAGAGGAGCCCGCCGCGCCAGAGCCAGCACTGGACAAGGGCTTGGGCACCG 60

Qy 7747 GCCAGGCTGCACTGTACAGGACCCCGGGAAGCCCTCTGGGACTGTCTGCTGGCAACGCC 7806

Db 61 GCCAGGCGCAGTGTACAGGACCCCGGGAAGCCCTCTGGGACTGTCTGCTGGCAACGCC 120

Qy 7807 GAGGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCACACCGCCGG 7866

Db 121 GAGGCCCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGCAGCATCACACCGCCGG 180

Qy 7867 AGGTCCAGTCTCACTTTCCAGAGTGAAGGGCATGAAGGAGCTGTCTGGTGG 7926

Db 181 AGGTCCAGTCTCACTTTCCAGAGTGAAGGGCATGAAGGAGCTGTCTGGTGG 240

Qy 7927 CCACCAAGATCAACTCGAGCGCCATCAAGTGCACCTACGGCACAGTGCAGAGTGCAGA 7986

Db 241 CCACCAAGATCAACTCGAGCGCCATCAAGTGCACCTACGGCACAGTGCAGAGTGCAGA 300

Qy 7987 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGAATCACTCTCTTTCTCAAGCGGACG 8046

Db 301 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGAATCACTCTCTTTCTCAAGCGGACG 360

Qy 8047 GCAAGGCTCTGAACCTACTGGGACTTTCGGACCGCTTGTGGGACCCAGCTCCG-CCT 8105

Db 361 CAAAGGCTCTGAACCTACTGGGACTTTCGGACCGCTTGTGGGACCCAGCTCCG-CCT 420

Qy 8106 TAGTCCCGCAACTCTGAGCCCATGTCTGCCCGCCAGCCCAAGGGGACAGGCTCACCTC 8165

Db 421 TAGTCCCGCAACTCTGAGCCCATGTCTGCCCGCCAGCCCAAGGGGACAGGCTCACCTC 480

Qy 8166 TACCAAAACCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGCATTTTCTCC 8225

Db 481 TACCAAAACCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGCATTTTCTCC 540

Qy 8226 AGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTAGAAATATAGTCTCTTCTCCAGAAATCC 8285

Db 541 AGCTCAGAACCCAGGGCTCTGCCCGAGTCTGTAGAAATATAGTCTCTTCTCCAGAAATCC 600

Qy 8286 CAGCGGCGCAATGGAACCTCAGCTGGGTCTTAATACAGTCTTTAAAGGCCCGAGCC 8345

Db 601 CAGCGGCGCAATGGAACCTCAGCTGGGTCTTAATACAGTCTTTAAAGGCCCGAGCC 660

Qy 8346 CTAGAAACCCAGCTCTCTCCCTCGG-AACCGCTCAGCTAGAGCCAGACCAAGCTTACTCAG 8404

Db 561 CTAGAAACCCAGCTCTCTCCCTCGG-AACCGCTCAGCTAGAGCCAGACCAAGCTTACTCAG 720

Qy 8405 GGCTCTCTCCAGTTGTAGGAGCTGAGGTTTACCGCTTTAACCCCAAGGGGACAGGCTCC 8464

Db 721 GGCTCTCTCCAGTTGTAGGAGCTGAGGTTTACCGCTTTAACCCCAAGGGGACAGGCTCC 780

Qy 8465 ACCTCCAGCCCGGG 8478

Db 781 ANCTTCAGCCCGGG 794

RESULT 7

US-09-726-211-1570

; Sequence 1570, Application US/09726211

; GENERAL INFORMATION:

; APPLICANT: Hunter, John J.

; APPLICANT: Shvjan, Andrew W.

; TITLE OF INVENTION: NOVEL NUCLEIC ACID MOLECULES AND USES

; IMMEDIATE SOURCE: THEREFOR

Query Match	7.7%	Score 706.2;	DB 29;	Length 784;
Best Local Similarity	98.4%;	Pred. No. 3.8e-47;		
Matches 724; Conservative	0;	Mismatches 9;	Indels 3;	Gaps 1;

QY	2261	GCAGGCGCGCCACACGCCGCCGCGCATGGGTGAAGCGCGCCAGAGCAGCTGTTC	2320
Db	47	ggaggccgcgcaccacgcgccgcgcatgggtgaagcgcgcagaagactgttc	106
QY	2321	CCGCTGGCGCTCATCGACGAGTGGTGCGCTGTTGCTGCGAGCTGGCGGAGAGGAG	2380
Db	107	ccgctgcgctccatcgacgacgtgtagtgcgctgttctgcgcgagtggccgagagag	166
QY	2381	CCGAGACTGGTGCTCTTCTTGGTGCTGGGCTTGGGAGCAATTTCTGGCTGTCAAC	2440
Db	167	ccggacctggtgctccctttcccttggtggtgcttgtggagcaattttctgtggcaac	226
QY	2441	CGGCTCATCCCTACCACAGTTTCCGAGCTCACCTTCACAGC - --CAGCCCCGCCCCGAC	2497
Db	227	cgegcatccctaaccaacgttcccagctcccttcagccgaacccccgcgcccccac	286
QY	2498	CCGCTCGGCGCCTCACCTACTTTCCGTGGCGCCGACCTGTCTATCATCGCGCCTCTAT	2557
Db	287	ccgcctggcgctccactactttcccgtgcggaacctgtcatacgcgcgcctctat	346
QY	2558	GCCGCTTCACCGCCCAGATCCGAGGCGCGTCGACCTGTCCTTATCTTCGAGAAGGG	2617
Db	347	gcccgtctccgcgccagatccggagcgcgctgacgtgtccctatatctctggagaggg	406
QY	2618	GGTGTCTCCAGCCGTGAGCTGGTGAAGAAGGTCTCCGATGTCTATGAACAGCCTCAGC	2677
Db	407	ggtgtccagccgtgagctggtgaagaaggtctccgatgtcataatgaacagacctcagc	466
QY	2678	CGTCTCTACTTCAAGGATCGGCGCCACATCCAGTCCCTCTTCAGTTCATCACAGTTGG	2737
Db	467	cgtccctacttcaaagatcgggccacatccagtcacctcttcagcttatcacaggttgg	526
QY	2738	AGCCACGTAGTGGGAATCTTATCCATCACCCACTTCTTCAAACCCCTCCATGTTTACA	2797
Db	527	agcccagttagtgggaattcttatccaacyaccacctctcacaacccctccatggttaca	586
QY	2798	GAACCCCTTTTAAGAACCTGTAAAGCTGTGAGGTTTCGGCAGCTGTATTCTCTTTGCA	2857
Db	587	gaaccttttaagaacctgaagccttgtgaggttcggcaggtgtatttctcttttga	646
QY	2858	GTTGGGAACCTGAAGCCCAGAGAGGGAATGATATGCCAAGTTCACACCGGATGGCA	2917
Db	647	gttgggaacctgaagccacagaggggaatgatatgccaaagtcacacagggcatggca	706
QY	2918	GGGCTGGAAGTGAAGCCTGTATCTGGCTCCAATCATCAACTTCACCTCTGCCCCCTC	2977
Db	707	ggcctggaagtgaagcctgatcaattggctccaatcatcaacctcaaccttgccccctc	766
QY	2978	AGCACCCCCACCCCTG 2993	
Db	767	aqacccccaccntg 782	


```
QY 2251 GGCTTGCTTGCAGCGCGCCGCCACCGCCGCCGCGGCTGAGGCGGCGCCAGAA 2310
Db 20 GAGTGTGGTGGTGGCGCGCGCCGCCACCGCGCGGCGGCTGAGGCGGCGCCAGAA 79
QY 2311 GAGCGTGTCCCGCTGGCTCCATCGACGAGTGGTGGCGGCTGTTGCTGCGGAGCTGGG 2370
Db 80 GAGCGTGTCCCGCTGGCTCCATCGACGAGTGGTGGCGGCTGTTGCTGCGGAGCTGGG 139
QY 2371 CCGAGGAGCGCGGACCTGGTGTCTCTTCTTGGTGTCTGGGCTTCTGGAGCATTTTCT 2430
Db 140 CCGAGGAGCGCGGACCTGGTGTCTCTTCTTGGTGTCTGGGCTTCTGGAGCATTTTCT 199
QY 2431 GGCTGTCAACGGGTCTATCCCTACCAAGTTCGGAGCTCACCTTCCAGCCGCGCGC 2490
Db 200 GGCTGTCAACGGGTCTATCCCTACCAAGTTCGGAGCTCACCTTCCAGCCGCGCGC 259
QY 2491 CCGGACCGCGCTGGCGGCTCACCTACTTCCCGTGGCGAGCTGCTATCATCGCGC 2550
Db 260 CCGGACCGCGCTGGCGGCTCACCTACTTCCCGTGGCGAGCTGCTATCATCGCGC 319
QY 2551 CCTCTATGCCCGCTTCCACCGCCAGATCCGAGGCGCGCTCGACCTGTCCCTCTATCCTCG 2610
Db 320 CCTCTATGCCCGCTTCCACCGCCAGATCCGAGGCGCGCTCGACCTGTCCCTCTATCCTCG 379
QY 2611 AGAAGGGGGTGTCTCCAGCGGTGAGTGGTGAAGAAGTCTCCGATGTATATGGAACAG 2670
Db 380 AGAAGGGGGTGTCTCCAGCGGTGAGTGGTGAAGAAGTCTCCGATGTATATGGAACAG 439
QY 2671 CCTCAGCGCTCTACTTCAAGATCGGGCCACATCCAGTCCCTCTTCAAGTTCATCAC 2730
Db 440 CCTCAGCGCTCTACTTCAAGATCGGGCCACATCCAGTCCCTCTTCAAGTTCATCAC 499
QY 2731 AGG 2733
Db 500 AGG 502

RESULT 15
PCT-US01-00663-5551/c
; Sequence 5551, Application PC/TUS0100663
; GENERAL INFORMATION:
; APPLICANT: Molecular Dynamics, Inc.
; APPLICANT: Penn, Sharon G.
; APPLICANT: Hanzel, David R.
; APPLICANT: Chen, Wensheng
; TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
; FILE REFERENCE: EB 0004 WO 7
; CURRENT APPLICATION NUMBER: PCT/US01/00663
; CURRENT FILING DATE: 2001-01-30
; PRIOR FILING DATE: US 60/180,312
; PRIOR FILING DATE: 04 February 2000 (04.02.00)
; PRIOR APPLICATION NUMBER: US 60/207,456
; PRIOR FILING DATE: 26 May 2000 (26.05.00)
; PRIOR APPLICATION NUMBER: US 09/632,366
; PRIOR FILING DATE: 03 August 2000 (03.08.00)
; PRIOR APPLICATION NUMBER: GB 24263.6
; PRIOR FILING DATE: 03 October 2000 (03.10.00)
; PRIOR APPLICATION NUMBER: US 60/236,359
; PRIOR FILING DATE: 27 September 2000 (27.09.00)
; PRIOR APPLICATION NUMBER: US 60/234,687
; PRIOR FILING DATE: 21 September 2000 (21.09.00)
; PRIOR APPLICATION NUMBER: US 09/608,408
; PRIOR FILING DATE: 30 June 2000 (30.06.00)
; NUMBER OF SEQ ID NOS: 38837
; SOFTWARE: Molecular Dynamics Sequence Listing Engine
; SEQ ID NO 5551
; LENGTH: 464
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: MAP TO AC000134.14
```

```
; OTHER INFORMATION: EXPRESSED IN PLACENTA, SIGNAL = 1.6
PCT-US01-00663-5551

Query Match 4.9%; Score 452; DB 1; Length 464;
Best Local Similarity 99.8%; Pred. No. 3.5e-27;
Matches 463; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 7714 AGAAGCCAGCACTGGACAAGGGCTGGGCACCGCGCAGGGTGCAGTGTCTCAGGACCCGCC 7773
Db 464 AGAAGCCAGCACTGGACAAGGGCTGGGCACCGCGCAGGGTGCAGTGTCTCAGGACCCGCC 405
QY 7774 GGAAGCCCTCTGGGACTGTCTGCTGGCACAGCCCGAGGCCCTTGAAGGTGGCAGCAGGCTC 7833
Db 404 GGAAGCCCTCTGGGACTGTCTGCTGGCACAGCCCGAGGCCCTTGAAGGTGGCAGCAGGCTC 345
QY 7834 AGTGCCAGCACCCCGCAGCATCACCGCGGAGGGTCCAGTGTCTACTTTCAGAGTG 7893
Db 344 AGTGCCAGCACCCCGCAGCATCACCGCGGAGGGTCCAGTGTCTACTTTCAGAGTG 285
QY 7894 AGAAGATGAAGGGCATGAAGGAGCTGCTGGTGCCACCAAGATCAACTCAGCGCCATCA 7953
Db 284 AGAAGATGAAGGGCATGAAGGAGCTGCTGGTGCCACCAAGATCAACTCAGCGCCATCA 225
QY 7954 AGCTGCAACTCACGGCACAGTTCGCAAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTA 8013
Db 224 AGCTGCAACTCACGGCACAGTTCGCAAGTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTA 165
QY 8014 GTGACTACACTGTCTTCTTCTCAAGCGGCAGCGCAAGGCTCTGAACTACTGGGACT 8073
Db 164 GTGACTACACTGTCTTCTTCTTCTCAAGCGGCAGCGCAAGGCTCTGAACTACTGGGACT 105
QY 8074 TCGGACCGCTTGTGGGACCCAGGCTCCG-CCTTAGTCCCGCAACTCTTGAGCCCATGTT 8132
Db 104 TCGGACCGCTTGTGGGACCCAGGCTCCG-CCTTAGTCCCGCAACTCTTGAGCCCATGTT 45
QY 8133 TGCCCCCAGCCCCAAAGGGGAGAGCGCTCACCTCTACCCAAACCC 8176
Db 44 TGCCCCCAGCCCCAAAGGGGAGAGCGCTCACCTCTACCCAAACCC 1

Search completed: October 14, 2001, 07:44:28
Job time: 37994 sec
```

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 03:51:57 ; Search time 218.79 Seconds
(without alignments)
7943.124 Million cell updates/sec

Title: US-09-380-337-3
Perfect score: 9180
Sequence: 1 CTGGTCTGAACTCTCTGGCC.....AGCCACGGGGCGCGCCGCCG 9180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued_Patents_NA.*

- 1: /cgn2.6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2.6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2.6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2.6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2.6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2.6/ptodata/1/ina/Backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	749.4	8.2	2196	US-08-865-337A-2	Sequence 2, Appli
2	746.4	8.1	812	US-08-865-337A-5	Sequence 5, Appli
3	471.4	5.1	538	US-08-865-337A-3	Sequence 3, Appli
4	348	3.8	14796	US-08-975-080-35	Sequence 35, Appl
5	347.6	3.8	31571	US-08-323-443B-1	Sequence 1, Appli
6	347.6	3.8	53526	US-08-658-136-2	Sequence 2, Appli
7	347.6	3.8	53577	US-08-658-136-1	Sequence 1, Appli
8	342.4	3.7	14796	US-08-975-080-35	Sequence 35, Appl
9	329.8	3.6	4421	US-08-257-963B-9	Sequence 9, Appli
10	329.8	3.6	4421	PCT-US95-07201-9	Sequence 9, Appli
11	325.2	3.5	17949	US-09-087-465-3	Sequence 3, Appli
12	324.8	3.5	3373	US-08-273-411-2	Sequence 2, Appli
13	324.8	3.5	8174	US-07-914-281-5	Sequence 5, Appli
14	324.8	3.5	8174	US-08-393-246-5	Sequence 5, Appli
15	324.8	3.5	8174	US-08-525-058A-5	Sequence 5, Appli
16	324.8	3.5	8174	US-08-696-731-5	Sequence 5, Appli
17	324.8	3.5	8174	US-09-042-531-5	Sequence 5, Appli
18	324.8	3.5	8174	PCT-US91-00899-3	Sequence 3, Appli
19	320	3.5	14636	US-09-173-914-6	Sequence 6, Appli
20	318.4	3.5	11613	US-08-484-044-10	Sequence 10, Appl
21	318.2	3.5	5543	US-08-687-080-101	Sequence 101, App
22	315.2	3.4	35060	US-08-814-095-7	Sequence 7, Appli
23	304.4	3.3	4803	US-09-197-636-1	Sequence 1, Appli
24	304.4	3.3	4803	US-09-197-636-3	Sequence 3, Appli
25	303.2	3.3	3286	US-09-211-417-2	Sequence 2, Appli
26	302.2	3.3	7620	US-07-767-135-1	Sequence 1, Appli
27	302.2	3.3	7620	US-07-841-652-1	Sequence 1, Appli

c 28	295.6	3.2	10380	4	US-09-077-354B-3	Sequence 3, Appli
c 29	294.8	3.2	14636	4	US-09-173-914-6	Sequence 6, Appli
c 30	293.2	3.2	1442	2	US-08-454-557C-120	Sequence 120, App
c 31	293.2	3.2	1442	2	US-08-340-426D-120	Sequence 120, App
c 32	293.2	3.2	1442	2	US-08-450-673C-120	Sequence 120, App
c 33	292.8	3.2	4421	2	US-08-257-963B-9	Sequence 9, Appli
c 34	292.8	3.2	4421	5	PCT-US95-07201-9	Sequence 9, Appli
c 35	291.8	3.2	3035	1	US-08-726-725-2	Sequence 2, Appli
c 36	290.4	3.2	22481	5	PCT-US95-07201-43	Sequence 43, Appl
c 37	290.4	3.2	35060	3	US-08-814-095-7	Sequence 7, Appli
c 38	290	3.2	4803	4	US-09-197-636-1	Sequence 1, Appli
c 39	290	3.2	4803	4	US-09-197-636-3	Sequence 3, Appli
c 40	289.2	3.2	72928	3	US-09-009-913-1	Sequence 1, Appli
c 41	288.2	3.1	3035	1	US-08-726-725-2	Sequence 2, Appli
c 42	288.2	3.1	6769	1	US-08-480-784-20	Sequence 20, Appl
c 43	288.2	3.1	6769	1	US-08-483-553-20	Sequence 20, Appl
c 44	288.2	3.1	6769	1	US-08-487-002-20	Sequence 20, Appl
c 45	288.2	3.1	6769	1	US-08-483-554B-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-08-865-337A-2
; Sequence 2, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2196 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT14
; CLONE: Consensus
; US-08-865-337A-2

```
Query Match      8.2%  Score 749.4;  DB 2;  Length 2196;
Best Local Similarity 97.1%;  Pred. No. 5.2e-148;
Matches 775;  Conservative 4;  Mismatches 17;  Indels 2;  Gaps 2;

Qy 7683 GAGTCCAAAGCCAGAGAGCCCGCCCGCCCAAGAGCAGCAGCTGGACAAAGGGCGCTGGC 7742
Db 1381 GAGNCCAAAGCCAGAGAGCCCGCCCGCCCAAGAGCAGCAGCTGGACAAAGGGCGCTGGC 1440

Qy 7743 ACCGGCCAGGGTCAGTCTCAGGACCCCGCCCGAGGCTCTGGAGCTGTGCTGGCACA 7802
Db 1441 ACCGGCCAGGGTCAGTCTCAGGACCCCGCCCGAGGCTCTGGAGCTGTGCTGGCACA 1500

Qy 7803 GCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGTGGCAGCAGCCCGCAGCATCACCCACG 7862
Db 1501 GCCGAGGCCCTGAAGTGGCAGCAGCGCTCAGTGGCAGCAGCCCGCAGCATCACCCACG 1560

Qy 7863 CCGAGGGTCCAGTCTCAGTCTTCCAGAGTGAAGATGAAGGCATGAAGAGCTGCTG 7922
Db 1561 CCGAGGGTCCAGTCTCAGTCTTCCAGAGTGAAGATGAAGGCATGAAGAGCTGCTG 1620

Qy 7923 GTGGCCACCAAGATCACTCAGCGCCCATCAAGCTGCAACTCACGGCAGCTCCCAAGTG 7982
Db 1621 GTGGCCACCAAGATCACTCAGCGCCCATCAAGCTGCAACTCACGGCAGCTCCCAAGTG 1680

Qy 7983 CAGATGAAGAAAGCAAGTGTCCACCCCTAGTGACTACACTCTGTCTTCTTCTCAAGCG 8042
Db 1681 CAGATGAAGAAAGCAAGTGTCCACCCCTAGTGACTACACTCTGTCTTCTTCTCAAGCG 1740

Qy 8043 CAGCGCAAGGCCCTCTGAACCTCTGAGGAGCTTCCGAGCGCTTGTGGGACCCAGGCTCCG 8102
Db 1741 CAGCGCAAGGCCCTCTGAACCTCTGAGGAGCTTCCGAGCGCTTGTGGGACCCAGGCTCCG 1800

Qy 8103 -CCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCGCCCGCAGCCCAAGGGGACAGCGCTCA 8161
Db 1801 CCTTAGTCCCGCAACTCTGAGCCCATGTTCTGCGCCCGCAGCCCAAGGGGACAGCGCTCA 1860

Qy 8162 CCTCTACCCAAACCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTT 8221
Db 1861 CCTCTACCCAAACCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGATTTT 1920

Qy 8222 CTCAGCTCAGAAACCCAGGCTGTGCCCCAGTCTGAGATATAGTCTCTTCTCCAGCA 8281
Db 1921 CTCAGCTCAGAAACCCAGGCTGTGCCCCAGTCTGAGATATAGTCTCTTCTCCAGCA 1980

Qy 8282 ATCCAGCGCGCCCAATGAACCTCAGCGTGGTCCCTAAATTAACAGTCTTTAAAGGCCCA 8341
Db 1981 ATCCAGCGCGCCCAATGAACCTCAGCGTGGTCCCTAAATTAACAGTCTTTAAAGGCCCA 2040

Qy 8342 GCCCCTAGAAACCAAGCTCTCTCTCGG -AACCGCTCACCTAGAGCAGACCAACGTTAC 8400
Db 2041 GCCCCTAGAAACCAAGCTCTCTCTCGGNAACCGTACCTAGAGCAGACCAACGTTAY 2100

Qy 8401 TCAGGGCTCTCCAGCTGTGAGAGCTGAGGTTTCACCCCTTAACCCAAAGGAGCAGG 8460
Db 2101 TCAGGGCTCTCCAGCTGTGAGAGCTGAGGTTTNACCCCTTAACCCAAAGGAGGANGG 2160

Qy 8461 TCCCACTTCCAGCGCGGG 8478
Db 2161 TCCCACTTCCAGCGCGGG 2178
```

```
RESULT 2
US-08-865-337A-5
; Sequence 5, Application US/08865337A
; Patent No. 5972649
; GENERAL INFORMATION:
; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murty, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN
```

```
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/865,337A
FILING DATE: Herewith
CLASSIFICATION: 530
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0305 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 812 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: BLADTUT02
CLONE: 1313372
US-08-865-337A-5
```

```
Query Match      8.1%  Score 746.4;  DB 2;  Length 812;
Best Local Similarity 97.2%;  Pred. No. 1.1e-147;
Matches 772;  Conservative 4;  Mismatches 16;  Indels 2;  Gaps 2;

Qy 7687 CCAAGCCAGAGGAGCCCGCCCGCCCAAGAGCAGCAGCTGGACAAAGGGCGCTGGCACC 7746
Db 1 CCAAGCCAGAGGAGCCCGCCCGCCCAAGAGCAGCAGCTGGACAAAGGGCGCTGGCACC 60

Qy 7747 GCCAGGTGAGTCTCAGGACCCCGCCCGGAGGCTCTCTGGGACTGTCTGGGACAGCC 7806
Db 61 GCCAGGTGAGTCTCAGGACCCCGCCCGGAGGCTCTCTGGGACTGTCTGGGACAGCC 120

Qy 7807 GAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGACCCCGCAGCATCACCCAGCCGG 7866
Db 121 GAGGCCCTGAAGTGGCAGCAGCGCTCAGGTGCCAGACCCCGCAGCATCACCCAGCCGG 180

Qy 7867 AGGTCTCAGTCTCAGTCTTCCAGAGTGAAGAGTGAAGGAGTGAAGGAGTGAAGGAGT 7926
Db 181 AGGTCTCAGTCTCAGTCTTCCAGAGTGAAGAGTGAAGGAGTGAAGGAGTGAAGGAGT 240

Qy 7927 CCACCAAGATCACTCAGGCGCCATCAAGTGCACCTCACCGCAGCTCGGAAGTGACA 7986
Db 241 CCACCAAGATCACTCAGGCGCCATCAAGTGCACCTCACCGCAGCTCGGAAGTGACA 300

Qy 7987 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGACTACTCTCTTCTCTCAAGCGGAGC 8046
Db 301 TGAAGAAGCAGAAAGTGTCCACCCCTAGTGACTACTCTCTTCTCTCAAGCGGAGC 360

Qy 8047 GCAAGGCCCTCTGAACCTACTGGGAGCTTCGACCGCTTGTGGGACCCAGGCTCCG-CCT 8105
Db 361 CAAAGGCCCTCTGAACCTACTGGGAGCTTCGACCGCTTGTGGGACCCAGGCTCCGCGCT 420

Qy 8106 TAGTCCCGCAACTCTGAGCCCATGTTCTGCCCCCGCCAGCCCAAGGGGACAGCCCTCACCTC 8165
```

Db 421 TAGTCCCAACTCTGAGCCCATGTTCTGCCCCAGCCCAAGGGGACAGGCTCACCTC 480
QY 8166 TACCCAAACCTAGGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGGATTTCTCC 8225
Db 481 TACCCAAACCTAGGTCCCGGTCCCGAGTACAGTCTGTATCAAAACCCAGGATTTCTCC 540
QY 8226 AGCTCAGAACCCAGGCTCTGCCCGAGTCTGTAGATATAGGTCTCTTCCAGAAATCC 8285
Db 541 AGCTCAGAACCCAGGCTCTGCCCGAGTCTGTAGATATAGGTCTCTTCCAGAAATCC 600
QY 8286 CAGCCGGCAATGGAACCTCAGCTGGGTCTTAAATACAGTCTTTAAAGGCGCCAGCC 8345
Db 601 CAGCCGGCAATGGAACCTCAGCTGGGTCTTAAATACAGTCTTTAAAGGCGCCAGCC 660
QY 8346 CTAGAAACCAACCTCTCTCTCGG - AACGCTCACCTAGAGCCAGACCAACGTTACTCAG 8404
Db 661 CTAGAAACCAACCTCTCTCGGNAACCGTTCACCTAGAGCCAGACCAACGTTACTCAG 720
QY 8405 GGCTCTCCAGCTGTAGAGCTGAGGTTTCAACCTTAAACCAAGGAGCAGAGTCCC 8464
Db 721 GGCTCTCCAGCTGTAGAGCTGAGGTTTNAACCTTNAACCAAGGAGGAAAGGTCCC 780
QY 8465 ACCTCAGCCCGG 8478
Db 781 ANCTTCAAGCCCGG 794

RESULT 3

US-08-865-337A-3
; Sequence 3, Application US/08865337A
; Patent No. 5972649

GENERAL INFORMATION:

; APPLICANT: Au-Young, Janice
; APPLICANT: Covitz, Peter
; APPLICANT: Tang, Y. Tom
; APPLICANT: Murry, Lynn E.
; TITLE OF INVENTION: HUMAN MULTIPLE ENDOCRINE NEOPLASIA TYPE
; TITLE OF INVENTION: 1 PROTEIN
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304

COMPUTER READABLE FORM:

; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/865,337A
; FILING DATE: Herewith
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0305 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:

INFORMATION FOR SEQ ID NO: 3:

; SEQUENCE CHARACTERISTICS:
; LENGTH: 538 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear

; IMMEDIATE SOURCE:
; LIBRARY: BRAINOT14
; CLONE: 1596949
US-08-865-337A-3

Query Match 5.1%; Score 471.4; DB 2; Length 538;
Best Local Similarity 98.3%; Pred. No. 5.3e-90;
Matches 475; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 2251 GGCTTGCCTTGCAGGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCAGAA 2310
Db 20 GCCTAGTGTGGATGCGCGCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCCGCCAGAA 79
QY 2311 GAGCCTGTTCCCGCTCGGCTCCATCGACGAGTGGTGGCGCTGTTTGGCTCCCGAGCTGGG 2370
Db 80 GAGCCTGTTCCCGCTCGGCTCCATCGACGAGTGGTGGCGCTGTTTGGCTCCCGAGCTGGG 139
QY 2371 CCAGAGGAGCGCGACCTGGTGTCTTCTTCTTGGTGGCTTCTGGAGCATTTTCT 2430
Db 140 CCAGAGGAGCGCGACCTGGTGTCTTCTTCTTGGTGGCTTCTGGAGCATTTTCT 199
QY 2431 GGCTGTACACCGCGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGCCCGCCCGCC 2490
Db 200 GGCTGTACACCGCGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGCCCGCCCGCC 259
QY 2491 CCCCAGCCCGCTGGCGGCTCACCTACTTCCCGTGGCGGCTGTCTATCATCAGCGC 2550
Db 260 CCCCAGCCCGCTGGCGGCTCACCTACTTCCCGTGGCGGCTGTCTATCATCAGCGC 319
QY 2551 CTTCTATGCCCCCTTCAACCGCCCGACATCCAGGCGCGCTGCGACCTGTCCCTCTATCTCG 2610
Db 320 CTTCTATGCCCCCTTCAACCGCCCGACATCCAGGCGCGCTGCGACCTGTCCCTCTATCTCG 379
QY 2611 AGAAGGGGTGTCTCCAGCGGTGAGTGGTGAAGAGGTCTCCGATGTCTATGGAACAG 2670
Db 380 AGAAGGGGTGTCTCCAGCGGTGAGTGGTGAAGAGGTCTCCGATGTCTATGGAACAG 439
QY 2671 CTTCAAGGCTCTTCAAGGATCGGCGCCACATCCAGTCCCTCTTCAAGTTCATCAC 2730
Db 440 CTTCAAGGCTCTTCAAGGATCGGCGCCACATCCAGTCCCTCTTCAAGTTCATCAC 499

QY 2731 AGG 2733
Db 500 AGG 502

RESULT 4

US-08-975-080-35/c
; Sequence 35, Application US/08975080
; Patent No. 6245523

GENERAL INFORMATION:

; APPLICANT: Altieri, Dario C.
; TITLE OF INVENTION: SURVIVIN, A PROTEIN THAT INHIBITS
; TITLE OF INVENTION: CELLULAR APOPTOSIS, AND ITS MODULATION
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
; STREET: 1800 M Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5869

COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,080
; FILING DATE: 20-NOV-1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,435

```

; FILING DATE: 20-NOV-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Adler, Reid G.
; REGISTRATION NUMBER: 30,988
; REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-467-7000
; TELEFAX: 202-467-7176
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 14796 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-975-080-35

```

Query Match 3.8%; Score 348; DB 4; Length 14796;
Best Local Similarity 66.8%; Pred. NO. 1.4e-63;
Matches 624; Conservative 0; Mismatches 285; Indels 25; Gaps 8;

[illegible]

APPLICANT: CONNORS, TIMOTHY D
APPLICANT: DACKOWSKI, WILLIAM
APPLICANT: GERMINO, GREGORY
APPLICANT: QIAN, FENG
TITLE OF INVENTION: POLYCYSTIC KIDNEY DISEASE GENE
NUMBER OF SEQUENCES: 58
CORRESPONDENCE ADDRESS:
ADDRESSEE: GENZYME CORPORATION
STREET: ONE MOUNTAIN ROAD
CITY: FRAMINGHAM
STATE: MASSACHUSETTS
COUNTRY: USA
ZIP: 01701
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
FILING DATE: US/08/658.136
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: LASSEN, ELIZABETH
REGISTRATION NUMBER: 31,845
REFERENCE/DOCKET NUMBER: GEN4-17.8
TELECOMMUNICATION INFORMATION:
TELEPHONE: 508-872-8400
TELEFAX: 508-872-5415
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 53577 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-658-136-1

Query Match 3.8%; Score 347.6; DB 3; Length 53577;
Best Local Similarity 78.7%; Pred. No. 2.8e-63;
Matches 496; Conservative 0; Mismatches 114; Indels 20; Gaps 6;

QY 3115 TTTTCTGATCGGCTCACTGACAGGCTCTCACTCTGCTCCCGGCTGAGTGCAGT 3174
Db 7496 TCTTTTGTGTTGTTGAGAGGAGTCTGGCTCTGTCAACCCAGGCTGGAGGACAA 7555

QY 3175 GGTGTGATCATGGCTCACTGACGCTCAGTCTCCCTGGGCTCAGAGATCCTCCACCTCA 3234
Db 7556 GGTGTGATCTCGGCTCACTGACGCTCCTATCTCCCGGTTCAAGCGATCTCCTGCCTCA 7615

QY 3235 GCCTCTGATGAGTACAGTACAGGATGACAC--CGTGGTGTCTGTTAAATTTTGTAT 3292
Db 7616 GCCTCTGATGAGTACAGTACAGGATGAGAGCGCGCCACACAGCGCGCTAATTTTAAAT 7675

QY 3293 TTTTCTGATGAGTGTCTCACTATGTGGCCAGACTGTCTTGAACCTTTGGGCTC 3352
Db 7676 ATTTTGTAGAGATGGGCTTTTCAACATGTTGGTCAAGCTGTCTTGAACCTTTGGCCTT 7735

QY 3353 AGTGTATGCGCGGCTCAGTCTCCCAA--TGCTGGGATTACAGGTGTAG--CCACC 3407
Db 7736 AGTGTATGCTGTGCTCGGCTCCCAAAGTCTCGGGATTACAGGTGTAGTGTAT 7795

QY 3408 GCAACTGTGGCTATGAAATTTTTTTTTTTTTCAGACGGCGTCTCACTCTGCGGCC 3467
Db 7796 TTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 7855

QY 3468 AGGCTGGAGTGCAGTGTGCAATCTCGGCTCACTGCAAGCTCTGCTCTGCTTTTCATGC 3527
Db 7856 AGGCTGGAGTGCAGTGTGCAATCTCGGCTCACTGCAAGCTCTGCTCTGCTTTTCATGC 7915

QY 3528 CATTCTCCTGCTCTGCTCAGCTCTGAGTACTGGGACTACAGGCTGCGCCACA 3587
Db 7915 CATTCTCCTGCTCTGCTCAGCTCTGAGTACTGGGACTACAGGCTGCGCCACA 3587

Db 7916 CATT-----CTCCTGCTCAGCCTCTCTGAGTAGCTGGACT---GGTGGCCGTCACCA 7965

QY 3588 TCCGTGCTAAATTTTTTTTGGATTTTGTAGTAGACAGGTTTCAACCATGTTAGCCAGG 3647
Db 7966 TGCCAGCTAA--TTTTTGTATTTTGTAGTAGACGGGGTTTCAACCTGTTAGCCAGG 8022

QY 3648 ATGGTCTCGATCTCTGACCTCGTGATCCGCCGCTTGGCTCCCAAAGTGTGGGAT 3707
Db 8023 ATGGTCTCGATCTCTGACCTCGTGATCTCCGCCCTCAGCTCCCAAAGTGTGGGAT 8082

QY 3708 ACAGCGTGAGCCAGCCGACCTGGTCAAAA 3737
Db 8083 ACAGCGTTGAGCCAGCCGCTGTCTTTAAA 8112

RESULT 8
US-08-975-080-35
Sequence 35, Application US/08975080
Patent No. 6245523
GENERAL INFORMATION:
APPLICANT: Altieri, Dario C.
TITLE OF INVENTION: SURVIVIN A PROTEIN THAT INHIBITS
CELLULAR APOPTOSIS, AND ITS MODULATION
NUMBER OF SEQUENCES: 35
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORGAN, LEWIS & BOCKIUS LLP
STREET: 1800 M Street, N.W.
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036-5869
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
FILING DATE: 20-NOV-1997
APPLICATION NUMBER: US/08/975.080
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031.435
FILING DATE: 20-NOV-1996
ATTORNEY/AGENT INFORMATION:
NAME: Adler, Reid G.
REGISTRATION NUMBER: 30,988
REFERENCE/DOCKET NUMBER: 044574-5022-01-WO
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-467-7000
TELEFAX: 202-467-7176
INFORMATION FOR SEQ ID NO: 35:
SEQUENCE CHARACTERISTICS:
LENGTH: 14796 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-975-080-35

Query Match 3.7%; Score 342.4; DB 4; Length 14796;
Best Local Similarity 76.0%; Pred. No. 2e-62;
Matches 497; Conservative 0; Mismatches 131; Indels 26; Gaps 5;

QY 3118 TTTTCTGATGAGTGTCTCACTCTGCTCCCGGCTGAGTGCAGTGTGT 3177
Db 5671 TTTTCTGATGAGTGTCTCACTCTGCTCCCGGCTGAGTGCAGTGTGT 5730

QY 3178 GTGATCATGGTCACTGAGCCTCAGTCTCCCTGGGCTCAGAGATCCTCCAACTCAGCC 3237
Db 5731 CTGATCTCGGATCACTGCAACCTCGGCTCCTGGGCTCAAGTGAATTTGCTTTCAGCC 5790

QY 3238 TCCGTGAGTGTGAGATGACAGGCTGAGTGTGCTGTTTAAATTTTGTATTTT 3297
Db 5790 TCCGTGAGTGTGAGATGACAGGCTGAGTGTGCTGTTTAAATTTTGTATTTT 3297

Db 5791 TCCCAAGTAGCGGAGATTACAGGATGTCGCCACACACCAGGTAATTTTGTGA--TTTT 5848
QY 3298 TTGTAGAGATGGTGTCTCACTATGTGGCCAGAGCTGTGTGAACCTTTGGGCTCAAGTG 3357
Db 5849 TGGTAGAGACGAGGTTTCCACCATGCTGGCCAGGCTGTTTGAACCTCTGACCTCAGGTG 5908
QY 3358 ATCTGCCCGCTCAGTCTCCCAA--TGCTGGGATTACAGGTGTGAGCCACCGGAACCTGG- 3415
Db 5909 ATCCACCCGCTCAGCCTCCCAAAGTCTGAGATTATAGGTGTGAGCCACACACCTGGC 5968
QY 3416 -----TGGCCTATGAAATTTTTCAGACGCGCTCTCACTCTGT 3462
Db 5969 CTCAGGAAGTATTTTATTTTAAATTTATTTATTTAGATGGAGTCTGCTCTGT 6028
QY 3463 CGCCAGGCTGGAGTGGTGTGCAATCTCGGCTCACTGCAAGCTCTGCTCTCTGCTTT 3522
Db 6029 CGCCAGGCTAGAGTCAGCAGCGGGATCTCGGCTCACTGCAAGCTCCGCCCCCGCAGGTT 6088
QY 3523 CATGCCATTCCTGCTGCTCTGCTCAGCTCTGAGTCTGAGTCTGAGGACTACAGGAGCTGC 3582
Db 6089 CAAGCCATT-----CTCCTGCTCAGCTCTCGGAGTAGCTGGGACTACAGGCGCCGC 6141
QY 3583 CACCATCGCTGGCTAAATTTTGTGATTTTGTAGTAGAGACGAGGTTTCAACCATGTTAG 3642
Db 6142 CACCAACCCGGCTAA--TTTTTTGTATTTTGTAGTAGAGACGAGGTTTCAACCGTGTAG 6199
QY 3643 CCAGGATGGTCTCGATCTCCTGACCTCGTGTATCGCCGCTTGGCCTCCCAAAGTCTG 3702
Db 6200 CCAGGAGGCTTGTATCTCTGACCTCGTGTATCTGCTGCTGCTGCTGCTCCCAAAGTCTG 6259
QY 3703 GGATTACAGCGGTGAGCCAGCGACCTGGTCAAAATGTTTGTAGACAGAGAAGG 3756
Db 6260 GGATTACAGGTGTGAGCCACACACCCGGCTATTTTATTTTGTAGACAGGG 6313

RESULT 9

US-08-257-963B-9
; Sequence 9, Application US/08257963B
; Patent No. 5840686
; GENERAL INFORMATION:
; APPLICANT: Chader, Gerald J.; Becerra, S.
; APPLICANT: Patricia; Schwartz, Joan P.;
; APPLICANT: Taniwaki, Takayuki
; TITLE OF INVENTION: PIGMENT EPITHELIUM
; TITLE OF INVENTION: DERIVED FACTOR: CHARACTERIZATION OF ITS NOVEL
; TITLE OF INVENTION: BIOLOGICAL ACTIVITY AND SEQUENCES ENCODING
; TITLE OF INVENTION: AND EXPRESSING THE PROTEIN
; NUMBER OF SEQUENCES: 42
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Morgan & Finnegan
; STREET: 345 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy Disk
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/257,963B
; FILING DATE:
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 07/952,796
; FILING DATE: 24-SEPT-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: DOROTHY R. AUTH
; REGISTRATION NUMBER: 36434
; REFERENCE/DOCKET NUMBER: 20264126US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4421 Base Pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: Genomic DNA
; ORIGINAL SOURCE:
; ORGANISM: Human
; FEATURE:
; NAME/KEY: J7101
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 7.1 kb Bam HI
; OTHER INFORMATION: fragment Derived from human placental
; OTHER INFORMATION: genomic DNA
US-08-257-963B-9

Query Match 3.6%; Score 329.8; DB 2; Length 4421;
Best Local Similarity 77.4%; Pred. No. 5.6e-60;
Matches 484; Conservative 0; Mismatches 117; Indels 24; Gaps 6;
QY 3117 TTTTTTTTTTTTTTTTTTTTGAGACAGGCTCTCACTCTGTCCCCCAGGCTGGAGTGCAGTGG 3176
Db 3491 TATTTTTTTTTTTTTTTTGTAGATGGAGTCTGGCTCTCTCACCCAGGCTGGAGTGCAGTGG 3550
QY 3177 TGTGATCATGGCTACACGAGCCCTCAGTCTCCCTGGGCTCAGAGATCTCCAACCTCAGC 3236
Db 3551 CGGATCTCGGCTCAGTCTGCAAGCTCTTCTCCCGGGTTTCACCCCCATCTCTCGCTCTCAAC 3610
QY 3237 CTCTGAGTAGTCTGAGAGTACAGGCTACAGGCTGGTGGTGTAAATTTTGTGATTTT 3296
Db 3611 CTCCGAGTAGTCTGTAATACAGGCGCTCCGCCACCGCCGACTAATTTTGTGATTT 3670
QY 3297 TTGTAGAGATGGTCTCTCACTATGTGGCCAGAGCTGGTCTTGAACCTCTTGGGCTCAAGT 3356
Db 3671 TTAGTAGAGACGGGGTTTACCGCTTGTAGCAGGATGGTCTGGATCTCTGACCTC--GT 3728
QY 3357 GATCTGCCCGCTCAGTCTCCCAA--TGCTGGGATTACAGGTGTGAGCCACCGCAACTGG 3415
Db 3729 GATCTGCTGCTCGGCTCCCAAAGTGTGGGATTACAGGCTTGAAGCTTGAAGCTTGAAGCTTGA 3788
QY 3416 TGGCCTATGAAATTTTTTTTTTTTTCAGACGCGCTCTCACTCTGTGCGCCAGGCTGGA 3475
Db 3789 CCTC-----TTATTTTTTTTTTTTGGATGGAGTCTCACACTGTACCTGGGCTGA 3839
QY 3476 GTGCAGTGGTCAATCTCGGCTCACTGCAAGCTCTGCTCTCTGCTTTTCATGCCATTTCTCC 3535
Db 3840 GTGCAGTGGAGCGATCTCGGCTCACTGCAAGCTCTGCTCTCTGCTTTTCATGCCATTTCTCC 3595
QY 3536 TGCCTCTGCTCAGCCTCTGCTGAGTCTGAGTCTGAGGAGCTGAGGAGCTGCCACCATGCTGC 3595
Db 3896 ---CTCTGCTCAGCCTCCCAAAGTGTGGGATTACAGGTTGAGGCTTGAAGCTTGAAGCTTGA 3952
QY 3596 TAATTTTTTTTTTTTGGATTTTGTAGTAGACGAGGTTTTCACCATGTTAGCAGGATGGTCTC 3655
Db 3953 TA---GTTTTTTTGTATTTTGTAGTAAGATGGGTTTTCACCATGTTTGGCCAGGCTGCTT 4009
QY 3656 GATCTCTGACCTC--GTGATCCCGCCGCTTGGCCTCCCAAAAGTGTGGGATTTACAGGC 3713
Db 4010 GAATCTCTGACATCAGGTGATCCGCCACCTTAGCCTCCCAAAAGTGTGGGATTTACAGGC 4069
QY 3714 GTGAGCCACCGCACTGTCACAAA 3738
Db 4070 GTGAGCCACCATACCTGGCCAGCAA 4094

RESULT 10

PCT-US95-07201-9
; Sequence 9, Application PC/TUS9507201
; GENERAL INFORMATION:

QY 3278 GTTAATTTTTTGTAT-TTTTTTTAGAGATGGTGTCTCACTATGTGGCCAGCACTGGTC 3336
Db 2938 AGCTAATTTTTTGTATTTTTTTAGTAGAGATGGGTTTTTGCCATGTGGCCAGCTGGTC 2879
QY 3337 TTGAACCTCTTGGCTCAAGTATCTGCCCGCTCAGTCTCCCAAA-TGCTGGGATTACAG 3395
Db 2878 TCAAGTCTGACCTCAGGTGATCCACCTACTTCGGACTCCCAAGTACTAGGATTACAG 2819
QY 3396 GTGTAGCCACCGCAACTGGTGGCTTATGAAATTTTTTTTTTTTTCAGACGGCTCTC 3455
Db 2818 GCATGAGCC-----ACTGTGGTCCCAACTGTATTTTTTTTTTTCAGATGGAGTCTC 2765
QY 3456 ACTCTCTGCCAGGCTGGAGTGCAGTGTGCAATCTCGGCTCACTGCAAGCTCTCCCTC 3515
Db 2764 GATCTCTTGGCCAGGCTGGAGTGCAGTGTGATTTTGGCTCACTGCAAGCTCCGCCCTC 2705
QY 3516 CTGCTTTTCATGCCATCTCCTGCCCTCCTGCCCTCAGCTCCTGAGTAGCTGGGACTACAGG 3575
Db 2704 CCGGGCTCATGCCATTT-----CTCCTGCCCTCAGCTCCTCCGAGTAGCTGGGACTATAGG 2652
QY 3576 AGCTGCCACCATGCTCGCTAA--TTTTTTTTTTTGGATTTTAGTAGAGACGAGTTTCA 3633
Db 2651 CTCGCCACCATGCCCGCTGATTTTTTTTTTTTGTATTTTAGTAGAGATGGGGTTTCA 2592
QY 3634 CCATGTTAGCCAGGATGGTCTCGATCTCCTGACCTCGTATCGGCCCGCTTGGCCCTCC 3693
Db 2591 CCGTGTGTAGCCAGGATGGTCTCGATCTCCTGAACCTGCTGACTCGCCTACCTCGGCCCTCC 2532
QY 3694 AAAGTCTGGGATTACAGCGGTGAGCCACCGCAC 3727
Db 2531 AAAGTCTGGGATTACAGCGGTGAGCCGATGCC 2498

RESULT 12

US-08-273-411-2
; Sequence 2, Application US/08273411
; Patent No. 5625124
; GENERAL INFORMATION:
; APPLICANT: Falk, Per
; APPLICANT: Gordon, Jeffrey I.
; TITLE OF INVENTION: Animal Model for Gastro-Intestinal
; TITLE OF INVENTION: Disease
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Patrea L. Pabst
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: USA
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/273,411
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: WU106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (404) 815-6508
; TELEFAX: (404) 815-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3373 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna

; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 104..1201
; OTHER INFORMATION: /note= "Nucleotides 104 through 1201 encode the GDP-L-fuc
; PUBLICATION INFORMATION:
; AUTHORS: Larsen, et al.
; JOURNAL: Proc. Nat'l Acad. Sci. USA
; VOLUME: 87
; PAGES: 6674-6678
; DATE: 1990
; RELEVANT RESIDUES IN SEQ ID NO: 2: FROM 1 TO 3373
US-08-273-411-2

Query Match 3.5%; Score 324.8; DB 1; Length 3373;
Best Local Similarity 76.0%; Pred. No. 5.6e-59;
Matches 491; Conservative 0; Mismatches 122; Indels 33; Gaps 6;

QY 3116 TTTTTTTTTTTTTTTTTTTGGAGACAGGGTCTCACTCTGTCCTCCCGCAGCTGAGTGCAGTG 3175
Db 1753 TTTTTTTTTTTTTTTTTTTGGAGACAGGGTCTCGCTCTGTTGCCCGCAGCTGGAGTGCAGTG 1812
QY 3176 GTGTGATCATGGTCACTCGAGCCTCAGTCTCCCTGGGCTCAGAGATCCCTCCCAACCTCAG 3235
Db 1813 GCCTGATCTGGCTCACTGCAACTTCCGCTCTCTGTTCAAGCATTCCTCTCTCTCAG 1872
QY 3236 CTTCTGATGATGAGTACAGGATGACCGTGGTGTGCTGTTAAATTTTTTTTGTATTTT 3295
Db 1873 CTTCTGATGATGAGTACAGGATGACCGTGGTGTGCTGTTAAATTTTTTTTGTATTTT 1930
QY 3296 TTTTGTAGAGATGGTCTCACTCTGTCCTGAGCTGCTGCTTGAACCTCTTGGGCTCAAG 3355
Db 1931 TTTAGTAGAGACAGGGTTTCCACCATGTTGGCCGGGATGGTCTCGATCTCTCGACCT--TG 1988
QY 3356 TGATCTGCCCGCTCAGTCTCTCCCAAA--TGCTGGGATTACAGGCTGTCAGCCACCGCAAC-- 3412
Db 1989 TCATCCACCTGCTTGGCCCTCCCAAGTCTGGGATTACTGGCATGAGCCACTGTGCCCA 2048
QY 3413 -----TGCTGGCCTATGAAATTTTTTTTTTTTTTTTTCAGACGGCTCTCAC 3457
Db 2049 GCCCGGATATTTTTTTTAAATTTATTTATTTATTTATTTATTTATTTAGACGGAGTCTTGC 2108
QY 3458 TCTGTGCCCGCAGCTGAGTGCAGTGGTCAATCTCGGCTCACTGCAAGCTCTGCCCTCCT 3517
Db 2109 TCTGTAGCCCGCAGAGTGCAGTGGCGGATCTCAGCTCACTGCAAGCTCTGCCCTCCC 2168
QY 3518 GCTTTTCATGCCATTTCTCCTGCCCTCCTCAGCTCCTCCTGAGTAGCTGGGACTACAGGAG 3577
Db 2169 GGGTTTCATGCCATTT-----CTGCCCTCAGCTCCTGAGTAGCTGGGACTACAGGAG 2218
QY 3578 CTTGCCACCATGCTGCTGCTAAATTTTTTTTTTTTGGATTTTTTAGTAGACGAGGTTTACCAT 3637
Db 2219 CCGGCCACCATGCTGCTGCTAA--TTTTTTTTTGTATTTTAGTAGACGAGGTTTTCATCGT 2277
QY 3638 GTTAGCAGGATGGTCTCGATCTCCTGACCTCTGATCCCGCCCTTGGCCCTCCCAAG 3697
Db 2278 GTTAACACGAGATGGTCTCGATCTCCTGACCTCTGATCTGCCACCTCTGCCCTCCCAAG 2337
QY 3698 TGCTGGGATTACAGGGGTGAGCCACCGCAGCTGCTGCTCAAAATGTTT 3743
Db 2338 TGCTGGGATTACCGCGGTGAGCCACCATGCTGCTGCCCGGATAATTT 2383

RESULT 13
US-07-914-281-5
; Sequence 5, Application US/07914281
; Patent No. 5324663
; GENERAL INFORMATION:
; APPLICANT: LOWE, JOHN B.
; TITLE OF INVENTION: METHODS AND PRODUCTS FOR THE SYNTHESIS
; OF OLIGOSACCHARIDE STRUCTURES ON GLYCOPROTEINS,

Query Match	3.5%	Score 324.8	DB 1	Length 8174
Best Local Similarity	76.0%	Pred. No. 7.9e-59		
Matches 491	Conservative 0	Mismatches 122	Indels 33	Gaps 6
QY	3116	TTTTTTTTTTTTTTTTTTTGAGACAGGGTCTCACTGTGTCCCCAGGCTGGAGTGCAGTG	3175	
Db	6335	TTTTTTTTTTTTTTTTTTTGAGACAGGGTCTCGCTGTGTCCCCAGGCTGGAGTGCAGTG	6394	
QY	3176	GTGTGATCATGGCTCACTGACGCCTCACTGTCTCCCTGGGCTCAGAGTCTCCCAACCTCAG	3235	
Db	6395	GGTGTGATCTTGGCTCACTGCAACTTCCGCCTCCTGTGTTCGAAGCATCTCTCTGTCTCAG	6454	
QY	3236	CTCTCTAGTAGTGCAGAGTACAGGCATGCACCGTGGTGGCTGTAAATTTTTTTGTATTTT	3295	
Db	6455	CTCTCTAGTAGATGGGACTACAGGCACAGGCCATATGCTCTGGCTAAATTTTTTTGTA--TT	6512	
QY	3296	TTTTGTAGAGATGGTGTCCTACTATGTGGCCACAGCTGGTCTTGAACTCCTGGGCTCAAG	3355	
Db	6513	TTTAGTAGACAGGGTTTCACCATGTGGCCGGATGGTCTCGATCCTCTGACCT--TG	6570	
QY	3356	TGATCTCCCGCCCTCAGTCTCTCCAAA--TGCTGGGATACAGGTGTGAGCCACCCGCAAC--	3412	
Db	6571	TCATCCACCTGCTCTTGCCCTCCCAAGTGTGGGATAC TGSCATGAGCCACTGTGQCCA	6630	
QY	3413	-----TGCTGGCCCTATGAAATTTTTTTTTTTTTTTTTTTTTCAGACGGCGTCTCAC	3457	
Db	6631	GCOCGGATATTTTTTTTTTAATTATTTATTTATTTATTTATTTATTTAGACGGAGTCTTGC	6690	
QY	3458	TCGTCCGCCAGGCTGGAGTGCAGTGGTGCAACTCTCGGCTCACGTCAAGCTCTGCCTCCT	3517	
Db	6691	TCGTAGCCAGGCCAGTGCAGTGGCGGATCTCAGCTCACGTCAAGCTCTGCCTCCC	6750	
QY	3518	GCTTTTCATGCAATTCTCTCTGCTCAGGCTCTCTGAGTAGCTGGGACTACAGGAG	3577	

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 04:05:21 ; Search time 644.85 Seconds
(without alignments)
8938.725 Million cell updates/sec

Title: US-09-380-337-3

Perfect score: 9180

Sequence: 1 CTGCTCTTGAACCTCTGCGC.....AGCCACGGCGCGCGCCGCCG 9180

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601:*
1: /SID88/gcgdata/geneseq/geneseq/NA1980.DAT:*
2: /SID88/gcgdata/geneseq/geneseq/NA1981.DAT:*
3: /SID88/gcgdata/geneseq/geneseq/NA1982.DAT:*
4: /SID88/gcgdata/geneseq/geneseq/NA1983.DAT:*
5: /SID88/gcgdata/geneseq/geneseq/NA1984.DAT:*
6: /SID88/gcgdata/geneseq/geneseq/NA1985.DAT:*
7: /SID88/gcgdata/geneseq/geneseq/NA1986.DAT:*
8: /SID88/gcgdata/geneseq/geneseq/NA1987.DAT:*
9: /SID88/gcgdata/geneseq/geneseq/NA1988.DAT:*
10: /SID88/gcgdata/geneseq/geneseq/NA1989.DAT:*
11: /SID88/gcgdata/geneseq/geneseq/NA1990.DAT:*
12: /SID88/gcgdata/geneseq/geneseq/NA1991.DAT:*
13: /SID88/gcgdata/geneseq/geneseq/NA1992.DAT:*
14: /SID88/gcgdata/geneseq/geneseq/NA1993.DAT:*
15: /SID88/gcgdata/geneseq/geneseq/NA1994.DAT:*
16: /SID88/gcgdata/geneseq/geneseq/NA1995.DAT:*
17: /SID88/gcgdata/geneseq/geneseq/NA1996.DAT:*
18: /SID88/gcgdata/geneseq/geneseq/NA1997.DAT:*
19: /SID88/gcgdata/geneseq/geneseq/NA1998.DAT:*
20: /SID88/gcgdata/geneseq/geneseq/NA1999.DAT:*
21: /SID88/gcgdata/geneseq/geneseq/NA2000.DAT:*
22: /SID88/gcgdata/geneseq/geneseq/NA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	9179	100.0	9180	19	AAV57416 Tumour suppressor
2	1312.2	14.3	2772	19	AAV57415 Tumour suppressor
3	749.4	8.2	2196	20	AAV80659 Human multiple end
4	746.4	8.1	812	20	AAV80662 Human multiple end
5	471.4	5.1	538	20	AAV80660 Human multiple end
6	352.4	3.8	119950	20	AAV90201 Human ves1 gene.
7	348	3.8	14796	19	AAV27941 Survivin gene. Ho
8	347.6	3.8	53526	19	AAV94101 Human PKD1 gene.
9	347.6	3.8	53577	17	AAV18551 Human polycystic k
10	347.6	3.8	53577	19	AAV94108 Human PKD1 locus b
11	345.8	3.8	66566	21	AAA53450 Human thioredoxin

C 12	344.8	3.8	122186	22	AAC89560	Human histone deac
C 13	343.6	3.7	26664	21	AAA60207	Human prostate can
C 14	342.8	3.7	51474	22	AAV97846	Human neuroblastom
C 15	342.4	3.7	14796	19	AAV27941	Survivin gene. Ho
C 16	341.4	3.7	26928	20	AAC32184	Human prothrombin
C 17	336.6	3.7	122186	22	AAC89560	Human histone deac
C 18	336.2	3.7	39198	22	AAV58067	Human polyamine-mo
C 19	335.4	3.7	1470	14	AAQ47355	Myotonic dystrophy
C 20	335	3.6	11967	22	AAV97863	Human neuroblastom
C 21	331.8	3.6	160552	22	AAV2697	Human glycosyl sul
C 22	329	3.6	66566	21	AAA53450	Human thioredoxin
C 23	328	3.6	1963	21	AAA51400	Chromosome 16g tum
C 24	326	3.6	18402	21	AAV3705	CYBRP fatty acid
C 25	325.2	3.5	17949	22	AAV89228	Human signal trans
C 26	325.2	3.5	18647	21	AAV21059	Human low adenosin
C 27	325.2	3.5	18648	21	AAA34937	Human adenosine re
C 28	325.2	3.5	72604	20	AAZ10752	Genomic sequence o
C 29	324.8	3.5	3373	18	AAV76768	Human alpha 1,2 fu
C 30	324.8	3.5	8174	12	AAQ13332	GDP-Fuc:beta-D-gal
C 31	324.8	3.5	8174	18	AAV61677	Human alpha(1,2)-f
C 32	323.6	3.5	160552	22	AAV2697	Human glycosyl sul
C 33	323.2	3.5	8174	15	AAQ56908	DNA encoding a gly
C 34	322.6	3.5	32367	19	AAV35820	Human SHOX (short
C 35	321.8	3.5	32367	19	AAV35620	*Human SHOX (short
C 36	321.2	3.5	2918	21	AAV73083	*Human MCL1 gene 5'
C 37	321.2	3.5	3204	21	AAV73082	Human melanocortin
C 38	321	3.5	11967	22	AAV97863	Human neuroblastom
C 39	320.2	3.5	15577	19	AAV35616	SHOX gene prelimin
C 40	320	3.5	14636	22	AAV31864	Human DNA-PKcs cod
C 41	319.2	3.5	121162	21	AAC66548	Human kinesin-like
C 42	318.6	3.5	15071	21	AAZ58411	hnRNP A2 genomic c
C 43	318.2	3.5	5543	18	AAV75284	Nucleotide sequenc
C 44	317.6	3.5	49999	20	AAZ23902	Human LOBO homolog
C 45	316.6	3.4	45546	20	AAV23520	Human kidney amino

ALIGNMENTS

RESULT 1	
AAV57416	
ID	AAV57416 standard; cDNA; 9180 BP.
XX	
AC	AAV57416;
XX	
DT	11-JAN-1999 (first entry)
XX	
DE	Tumour suppressor gene MEN1.
XX	
KW	Familial multiple endocrine neoplasia type 1; FMEN1; MEN1;
KW	menin; tumour suppressor gene; cancer; marker; diagnosis;
KW	gene therapy; human; ds.
XX	
OS	Homo sapiens.
XX	
FH	Key Location/Qualifiers
FT	CDS
FT	2288..8060
FT	/*tag= a
FT	/note= "contains introns"
FT	exon
FT	1680..1766
FT	/*tag= b
FT	/number= 1
FT	intron
FT	167..2264
FT	/*tag= c
FT	/number= 1
FT	exon
FT	2265..2732
FT	/*tag= d
FT	/number= 2
FT	intron
FT	2733..4296
FT	/*tag= e
FT	/number= 2
FT	exon
FT	4297..4505
FT	/*tag= f

QY	2941	CTTTGGCTCCAAATCATCAACCTCACCCTCTGCGCCCTCAGACACCCACCCCTTGCCACTGA	3000
Db	2941		
		cttggtccaaatcatcaaacctcaacctgccccctcagcaccgccaccttgccactga	3000
QY	3001	ACAGCTACAGGAGTTCTTAAGCATGATGACACAGAGGGCGGACAGAGATTTAGGGGGCAAGA	3060
Db	3001	acagctcacagggagttcttaagcatgagacacagagagggcgagcagaattcagggggcaaga	3060
		AGATGAAATTTGGGCTGCATTTTGAGGCAGCTTAACACAAATAATATGCTATGAAGATTTTTTT	3120
Db	3061	agatgaaattgggctgcatattgagcgagttaaacacaaataatgctatgaagattttttt	3120
		TTTTTTTTTTTTTTTGTAGACAGGGTCTCACTCTGTCTCCCGCCAGGCTGGAGTGCAGTGGTGTG	3180
Db	3121	tttttttttttttttagacaggggtctcaactgtctccccaggttggaagtcagtggtg	3180
		ATCATGGCTCACTGCAGCGCTCAGTCTCCCTGGGCTCAGAGATCTCTCAACCTCAGCGCTCC	3240
Db	3181	atcatggctcaactgcaagcctcaagtcctccctgggclcgagagatccctcaaacctcagcctcc	3240
		TGASTAGCTGAGAGTACAGGCATGCACCGTGGTGTGTTAAATTTTTTGTATTTTTTTTG	3300
Db	3241	tgagtagctgagagtacaggcatgcacgtggtgctggttaatttttttatattttttttg	3300
		TAGAGATGGTGTCTCACTATGTGGCCAGAGCTGGTCTTTGAACCTCTTGGGCTCAAGTGATC	3360
Db	3301	tagagatggtgtctcaactatgtggcccagactggtctcttgaacctctgggtcgaagtga	3360
		TGCGCGCCTCAGTCTCCCAAAATCTGTGGATTACAGGTGTGAGCCACCGCAACTGGTGGCC	3420
Db	3361	tgccgcctcagtcctccaaaaagtctgggagtacagagtgtagcaccacgcaacttgggtggcc	3420
		TATGAAAAATTTTTTTTTTTTTCAGACGGCGTCTCACTCTGTGCCCGAGGCTGGAGTGCA	3480
Db	3421	taTgaaatttttttttttttttttcagaagcggtctcaactctgtgccagctggagtgca	3480
		GTGGTCAATCTCGGCTCACTCGAAGCTCTGCCCTCCTTTTCATGCGCATCTCCTGCCCT	3540
Db	3481	gtggtgcaatccggctcaactgcaagctctgcctccctgtctctatgcaattccctgcct	3540
		CTGCGCTCAGCGCTCTGAGTAGCTGGGACTACAGAGCGCTGCCACCATGCGCTGGCTAAAT	3600
Db	3541	ctgctcagcctcctgtagtagctgggactacagagcctgccaccatgctggcctaatt	3600
		TTTTTTTTTGATTTTTTAGTAGACAGAGGTTTACACATGTTAGCCAGAGATGGTCTCAATCT	3660
Db	3601	tttttttgatttttagtagagacgaggtttcaccatgttagccagagatggtctcga	3660
		CTGACCTCTGTATCCGCCCGCTTTGGCTCTCCCAAGTCTGGGATTACAGGGCTGAGCC	3720
Db	3661		
		cttgacctctgctacgcgcgcgccttggtctcccaagtgctgggattcacagggcgtagcc	3720
QY	3721	ACCGCACCTGGTCAAAATTTGTTGACAGAGAGGGGCTTTGACCTCAAAAGCGTTAAGA	3780
		accgcaccitggtcacaaaatggttagacacagagagggccttgacctcaaaagccttaaga	3780
QY	3781	GTACAGGCTTGCANAAGAGCTTTGCACCAAGCCGGTTGACTGGCAATCCCATCTCTGGTGT	3840
		gtcagggcttgcaaaagagcttgacccaagccgggttgactgggcaatcccatcctcgtgt	3840
QY	3841	GCCATATTGAGAGGAATCAGAGGCTGCTTCTCAGCTTAGCAGCAAAAGAGTGCAGAGAT	3900
		gccaattatgagaagggaatcagaggtgtctctcagcttagcaggaaaagagtgcaagata	3900
QY	3901	AAATGAGGGTTATTTTGTGGTGGGTGTATAGCCAGAGAGTGTGGCCAGAGCTCTCTGTTTT	3960
		aaatgagggttattgttgtgtgggtgtagaccagagagtgtagccagcgtcctgtgttt	3960
QY	3961	TGCCATTCCTGTTTTTAACCTAGTAGTGAGTGAATGGAATTCCTTAATCCATAGAATA	4020
		tgcattcctgttttaacctagtagtgagtgaaatggaaatccctaaatacagagata	4020

Qy	4021	TATAATAGACTTCCAGAGAAACGACGAGTTAGGGCCAAAGGCTTGUSTCAGCTACAGGATAAT	4080
Db	4021	tataatagagtttcagagaaaacagaggttagggccaaaggctgggtcagctacaggata	
Qy	4081	CCAGAAAGGTATCTTTGGACATAGAGGTTAAACAGGAGAGAGTCTTTGAACACGT	4140
Db	4081	ccagaaagatactctgtgacatagaggggtataaacaggagagagctcttgaacaagct	4140
Qy	4141	GGAGGGAAGGATGAGGAGATAGTGGGCAGAGAAATCTGAGGTTGGTTCACAGGCTTGG	4200
Db	4141	ggaggaagggatggagggatagtgggcaggagaaactgaggttgggtcacagagctgg	4200
Qy	4201	AAAGGGAGTGGGAGGAGTGTGCCCATCACTACCTGTGCCCTTTCCCATGTGTTAAAGCA	4260
Db	4201	aaagggaagtgggagggagtgtyggccatacactacctgtgccctcttccccatgtlaaagca	4260
Qy	4261	CAGAGGACCTCTTTCACTACCTCCCTTCACAGGCACCAAAATGGACAGTCCGGTG	4320
Db	4261	cagagacctcttcatcactcccccttccaaaggcacaaaattggacagctccggtg	4320
Qy	4321	TGGCCTTGCTGTGGTGGGGCTGCCAGGCCCTGGGTCTCGGGATGTCACACCTCGCC	4380
Db	4321	tggccttgctgtggttggggcctccagggcctgggtctccggagtgccaactcgccc	4380
Qy	4381	TGCTCAGGATCATGCTGGGTAGTGTTTGGGCCCAATGGGGACACAGAGCTGAGGTCA	4440
Db	4381	tgctcagatcatgctcgtggtagtgtttgggcccatagggagcagagctgaagtca	4440
Qy	4441	CTTGGCAGCGCAAGGCAACGAGGACCGCAGGGGCCACAGTCAATGCTCGGTGTGGCTG	4500
Db	4441	cctggcacgaaaggcaacagagacggcaggggcccagagagtcacatgcccgtgtggtctg	4500
Qy	4501	AGCGGTTATTTGTTCCCTCCCTCCAGCTGTGCCCTTCATCTACTGTAGTGGCCAAAGCAC	4560
Db	4501	agcgggtattgttccctccccagccttgctcccttcatacttagtagcccaagccac	4560
Qy	4561	CCAAGGACTTCCATTTCTTGGGCCACACCCCTTTTCCCATCACACCACATAGGAA	4620
Db	4561	ccaaggacttccatttcttggggcacaccccttcttcccatcaccaccacacataggaa	4620
Qy	4621	GGGAAGACACAGAGCCCTTTTCTTGGCTGTCAATTCCTGAAGCAGGCACAGGTTGGC	4680
Db	4621	gggaagacagagagcccttctctggctgtcaatccctgaagcagcagcaggggtgggc	4680
Qy	4681	CATCATGAGACATAATGATCTCATCCCCCTTAAGAGTGGCTGTACCTGAAGAGTAT	4740
Db	4681	catcatgagacataatgatactcatcccccttaagagctggctgtcactgaaaggaatcat	4740
Qy	4741	ACATGCGCTGTACCCGACAGATGGAGTGGCCTTCATGGTGTGCCATCAACCTTCCA	4800
Db	4741	acatgctgtgaccgcaagaagggtggcctcatgggtgtgccaatcaaccttccca	4800
Qy	4801	TTGACTGTACACCGACTCGCTGGAGCTTCTGCAGCTGCAGAGGTGAGGGCTGAGCCAA	4860
Db	4801	tgaactgtcacacagactcgcgtggagcttctcagctgcagcaggtgagggctgaagccaa	4860
Qy	4861	TGGGGCAGGACTGGGCTAGCCAGACTTGACTTGCTGTGGAGACCTGGGCAGGGGCACTT	4920
Db	4861	tggggcagagactgggttaggccagacttgacttgcttgggacctgggcagggggcactt	4920
Qy	4921	TCCCTCTCAGCTTCAGTTCCCTTCCTGAAAAATGGTTAGTAAT	4980
Db	4921	tccctctcagcttcagcttcccccttccggaaaaatgggttagcaatctctggccttggc	4980
Qy	4981	CTTTCCCAGGGCTCTTTGGGAGATAGAATTGAGATGTGAAATTCCTTTGACTTCCATTTAAA	5040
Db	4981	ctttcccagggtcttgggagagtagaaattgagatgtgaaattgcttctgactcattaaa	5040
Qy	5041	GGGCTGTCCCAANAATTTGGCCCTTCCACATGCTGGGTGGTCCCTGTGGTCTGACCC	5100
Db	5041	gggctgtcccaanaatttggcccttccacatgggtgggtgggtccctgttggctctgaccc	5100
Qy	5101	CCACTCTGCCCATAGGCTTAAGGACCCGCTTCTCCTCCTGTTCGGTGGCTATTAAGTCT	5160

Db 5101 |||||ccccataggtcaagaccggtctcctccctgttccgtggtcatalaactct 5160
QY 5161 CTCTCTCGGCTCTAGAAAGTGTCTGGCTGCTTATGACCTGGGACATCTGGAAGGTC 5220
Db 5161 ctctctcgctctcagaagctgctcggtcgctctctcgaacctggaacctggaagtc 5220
QY 5221 AGTAGAGGAAGTGGCCAGGCTCGCCTGTGTGAGGCCGGGGGCTGGGTGGCAGCCTGAA 5280
Db 5221 agLagagggaagtggccagctcgctcggtgagcgcggtggtggtgagcctgaa 5280
QY 5281 TTATGATCCTTTCTTAGTACCCATGGCCCTTAGGGAAACCTGGAGATCTAGAGGAGCTG 5340
Db 5281 ttatgatcctttctaggtaccctatggccttaggaacctggcagatctagaggagctg 5340
QY 5341 GAGCCACCCCTCGCGCGCCAGACCACCTCACCTCTPACCACAGGPGGGGGCATCTAAG 5400
Db 5341 gagcccaacctggcgccagaccactcaacctctaccacaagggtgggggcatcgaag 5400
QY 5401 GAGGGTGCAGAGGAGACCTTAACAGTGGCTGAGGCAGGGGCCCTCATCTGGGCAGATG 5460
Db 5401 gaggggtgagaaaggagacctaacagtggctgaggcaggggcccctcatctgggcagatg 5460
QY 5461 AGAAGAGAACTTTGTGTGTGGGGGTATCGCCCATCCAGTCTCACTTTGTGTCAACTGT 5520
Db 5461 agaagagaactttgtgtgtgggggtatcgcccatccagtctcaactttgttcaactgt 5520
QY 5521 GTCCAGAACTCAGTTCAGTGAGGCTGTCTGAGGGGTGTCCAGGGTTCCTCCAGCCTGGGAG 5580
Db 5521 gtggaatacagtcacgtcagtcagggtgtctgaggggtgtccaggggtccccagcctggag 5580
QY 5581 TGCCAGGGGTGCTATTGTCCCTCAGCCCTGCTTTTCTGCCACTGCTTACTGTCTCTC 5640
Db 5581 tggcagggtgtcatttgtccccctcagcctgcttttctgcccactgttaactgtcccttc 5640
QY 5641 CTGGAGTATACAGAGGTCAATGTGGTAGGACACTGAAGAGGGGTGTTCATCTGGTG 5700
Db 5641 ctggagtataacagaggtcaaatgtgtagagcactgaagaggggtgttctacttggtg 5700
QY 5701 GGTGTAGGTGGGAGAGGCCAATTGGGCTGGCTGTGAAAGTCTTGGTATCTGTAGAA 5760
Db 5701 ggtgtagggtgggagagggccattgggtgggttgaaagtcttgggtgagtgtagaa 5760
QY 5761 GAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGGCGCAGGCCACCCCTCAGTCTG 5820
Db 5761 gagtgtctgagaaagagagggccctgagctcgagggcagggcccaacctgcagtctg 5820
QY 5821 CCCCAGCCTCAGCCAGCAGTCTGTAGACCCAGGGAGGAGACAGGTAGAAAGGCTGGC 5880
Db 5821 ccccaggcctcagcagcagctcctgtagaccaggagagaccaggttagaagggtggc 5880
QY 5881 AGCGAGTGGAGGTGGAGTGGAGAGTCCCTGGGATCTCCCTGGGATCTCCTGGCCCTT 5940
Db 5881 agcagtgaggggtggaggtggagagagactccctggggtctctcgtggccctt 5940
QY 5941 CTGGGTGTGCCCTGGTGGGSCATTTGTGCCAGAGGCGAGCTGGGGTGCCTCCCTGAGG 6000
Db 5941 ctgggtgtgccctgtggtgggcatttgtgccagcagggcagctgggtgcctccctgag 6000
QY 6001 ATCCTCTGCCTCAGCTTCATCCAGGCGATTCCTCAGCCAAAGACCTACTATCGGGATGAA 6060
Db 6001 atcctctgctcaacctccatccagggtcattgctctcagcaagaccctactactcggtgaa 6060
QY 6061 CACATCTACCCCTACATGTATCTTGGCTACCTGTCCGCAACCGCAATGTGCGGGAA 6120
Db 6061 cacatcacccctacatgttacctggctggctaccactgtgcgaacctggtgcgggaa 6120
QY 6121 GCCTGCAGCCCTGGCGGACAGCGCCACTGTATCATCAGAGGTGAGGATCCCCCTACTAG 6180
Db 6121 gccctcaggccttggcgacacggccactgtcatccaggagtggagatcccccactag 6180
QY 6181 GGCCCTGACGCTGTCTCTTCTCCCTTCCATCATGATTTCCAAACCCCTCGTCCAGGACTG 6240
|||||ccccatctgcaatgggagggcaggtcccaagcctgtgtgcaagtgggctgggc

Db 6181 ggctcgagcgtgtccctttctcctccctccatcagtttcccaaccacccctgttccaggactg 6240
QY 6241 AGGCCTCGCTCCACAGCCCGCATCCCTTTCCATCCAGTCCCTAGCAGCAGAGGCCACCAT 6300
Db 6241 aggcctggctcccaagcccccatccctttccatccagtccttaggcagaggaaccaat 6300
QY 6301 TACCAGAGAGGTAGGACCCCTGATTAAAGGTGTACATCTTTTCCCTCCCTCCCTCTCTCTC 6360
Db 6301 taccagagggtagggacctgattaaaggtgtcaactcttccctccctccctccctc 6360
QY 6361 CTAAATTTTTTTTTTCTCAGAACAGTCTCAAAATCTCCAATGTTTAAACCACCAATCATCAGC 6420
Db 6361 ctaattttttttctcagaacagttcaaatctccaatgttttaaccacatcatccagc 6420
QY 6421 AGTGGGACTTCCACCCCTCGGCCCATCCCTCCCTCTCATCTTTCTCTCTCTCTGGG 6480
Db 6421 agtgggacttcccaacctcgggcccatgtcccccctccctcatctctgtcttctcctc 6480
QY 6481 CTGACCCAGACAGCATCTTTTGAGTGAGTACCCACCTACTCTCCCGCAGCCCTGGGGG 6540
Db 6481 ctgacccagacagcatcattttgagtgagagacccaacctactccccagacccctgggg 6540
QY 6541 CTCCATCCCCCGCAGTCCCTTGGGGCTACCCCGATGGGTGAGACCCCTTTCAGACCCCTAC 6600
Db 6541 ctccatcccccgccaggtccctgggtgctacccccgatgggtgagaccccttcagaccctac 6600
QY 6601 AGAGACCCCACTGCTCTCAGACTACAACTACTTCCCGGGGAGAGCAGAGATCTACAAGG 6660
Db 6601 agagaccccaactgctctcagctacacactactgctcgccgggagagcagggagatctacaagg 6660
QY 6661 AGTTCTTTGAAGTAGCCAATGATGTCTATCCCAACCTGCTGAAGGAGCGACAGCTTGC 6720
Db 6661 agttctttgaagttagcaaatgagtgtcatccccaaacctgtctgaaggagggcagcagcttgc 6720
QY 6721 TGAGAGCGGGCGAGGAGCGCGCGGGGAGCAAGCCAGGTGAAAGGTGGAGCTTCCAGCC 6780
Db 6721 tggagggcgggagggagcgccgggggagcaagccaggtgaaagggctggaggtccagcc 6780
QY 6781 TGTGTCCAGCCTCCCACTGGACAGGCTCCCTTCCACAGGGCCATGGGGCTGCATGTA 6840
Db 6781 tgtgtccagctcccaacctggacaggggtcccttcccaagggccatgggggctgcatgta 6840
QY 6841 CGGGATTAGGAGTGGCAGGAGAGGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGTCT 6900
Db 6841 cgggattagggaatggcagggaggaagtggcccttgagcagacagctatgttcccttttgc 6900
QY 6901 ATAACTGAGTCTCTGGGCCCACTGGACGGACTGAAGGTATTTTATAGAGTTTCTACCC 6960
Db 6901 ataactgaggtcctggcccaagtggacgggagctgaaggtatttttagaggtttctacc 6960
QY 6961 TGTGCCCTTACATGATGCGCAGACTCCCTCCCTCAGCTGAGGGGTGAGGTAGGGATGG 7020
Db 6961 tgtgtccctcagtttcaatggccactccctccctcagctgaggggtggaggtaggatgg 7020
QY 7021 TAGTCTCTGCTATGATTTGGCTTTTATAAAGGAAAGAGGTCTTAAGAATGTTCCCAACC 7080
Db 7021 tagtctcctgctatgagattggctttataaaaggaagaggttctaaagaaagtctcccaacc 7080
QY 7081 TATGCTTACCTTTTCTGGAGCCAGGGGTCTTTGCTAGTGGGGGGCTTGGCTGTGGCC 7140
Db 7081 tatgtctacctttcttgagccaggggtctttgctcaggtgggggggctggcctgtgccc 7140
QY 7141 TCTGTCTAAGGGGTGAGTAAGAGACTGATCTGTGCCCTCCCTCCCTCCCTCCCTCCAGGGCAC 7200
Db 7141 tctgtctaaagggggtgagtaagagactgactgtgtccctccctccctccctccctccctcc 7200
QY 7201 CCAGAGCAAGGTTCGCCCTCCAGGACCTGAGTGTCTGCCCACTGCTGCTGCTGCTGCTGCT 7260
Db 7201 ccagagcaagaggttcgcgcctccagacacctgagtgctcgcaccaacctgtctgagatctc 7260
QY 7261 CGACGGCATCTGCAAAATGGGAGGAGGCGAGTCCACGCCCTGTGTGTCAGCTGGGCTGGGC 7320
Db 7261 cgacggcatctgcaaatgggagggagggcagttcccaagcctgtgtgtcaggtgggctgggc

Qy	7321	CACCTTTCTTGTCAGTCCCTTAGCGCGTTTTCAGGGACAGGTGAGGACAGCTGCACAGA	7380
Db	7321	caaccttcttgtagctccctaggccggttttgaggagacagtgaggagacagctgacacaga	7380
Qy	7381	GGTCTGGGCACATACAGGTGGTCAGACAGCCACGGCGTTGTACACATTTTCTTGCCCCAGG	7440
Db	7381	ggctgggcactacagtggtgacagcagcaaggctgtgcagacttttctggcccagg	7440
Qy	7441	GGCAGCATCTGCCCATCCCTTCGGTGGCGATGGGACATGACACCCCTGGTGGGATGGG	7500
Db	7441	ggcagcactctgcccctccctctcgggtgcgaTgggagactgagaccctccctgggtgggaTggg	7500
Qy	7501	ATGGCCAGACAGGGTCTCTGGAGTTCCAGCCACTGGCCGGCAACCTTGCTCTCACCTTGC	7560
Db	7501	atggccagagcagggtccctggagttccagccactggccgTcaaccttgctctcaecttgc	7560
Qy	7561	TCCTCCCACTGGCCAGAGTGGCGCAGAGAGTGGCCATAGTCAGCCGAGAGCGCGAGCGG	7620
Db	7561	tctcccactggcccaggTgcgcgcaagagtgcgcatagtgagccgagagggccgagggcgg	7620
Qy	7621	CCGAGGCCGAGGACCGTGGGGCGAGGAAGCCCGGAAGCGCGCGGGGGCCACGCGC	7680
Db	7621	ccgagggccgaggagccgtggggcggggaagccggggagggcgggcgggccacggc	7680
Qy	7681	GGAGGTCCAAAGCCAGAGGACCCCGCCGCCCAAGAACCCAGCACTGGCAAGGGCGCTGG	7740
Db	7681	gggagTccaaagccagagagccccgcgcgcccagaagcagcactggacaaagggccctgg	7740
Qy	7741	GCACCGGCACAGGTGCAGTGTGAGACGCCCGCCGGAAGCCCTCTGAGACATGTGCTGGCA	7800
Db	7741	gcaaccggcagggtgcagtgctcagtgacggccccgggaagcctcctgggactgtcgctggca	7800
Qy	7801	CAGCCGAGGCCCTGAAGGTGGCAGCAGCGGTCTAGTGCCAGCACCGCAGCATCACAC	7860
Db	7801	cagccggagggccctggaagTtggcagcaagggctcaggtggcagccaccgcagcaTcaaccac	7860
Qy	7861	CGCCGAGGGTCCAGTGCTCACTTTCCAGGTGAGAAGATGAAGGGCATGAAGGAGCTGC	7920
Db	7861	cgccggagggtccagTgctcactttccagagtgagaagatgaaagggcatgaaaggagctgc	7920
Qy	7921	TGTTGGCCACCAAGATCACTCGAGGGCCATCAAGCTGCACTCAGGGCACAGTGGCAAG	7980
Db	7921	tggtggccaccaagataTcaactcgagcgccatcaagctgcaactcacggcacagTcgcaag	7980
Qy	7981	TGCAGATGAAGAGCAGAAAGTCTCACCCCTAGTGACTACACTCTGTCTTCTCTCAAGC	8040
Db	7981	TgcagaTgaagagcagaagTgtccaccccttagTgactacactctgtctctctcctcagc	8040
Qy	8041	GGCAGCCAAAGGCCCTCAACTACTTGGGGACTTCGGACCGCTGTGGGACCCAGGCTC	8100
Db	8041	ggcagcgcaaaagccctctgaactactggggacttcgaccgcttgTggggaccacaggtc	8100
Qy	8101	CGCCTTAGTCCCCAACTCTGAGCCCATGTCTGTCCCCCAGCCCAAGGGGACAGGCTC	8160
Db	8101	cgcttagtcccccaactctgagccatgttctgccccagcccaaggggacaggtcct	8160
Qy	8161	ACCTCTACCCAAACCTAGTGTCCCGGTCCCGAGTACAGTCTGTATCAAACCCAGGATTT	8220
Db	8161	acctctaccacaaccctaggttcccggtcccgagTaccagctgtatcaaacccccagatt	8220
Qy	8221	TCTCCAGCTCAGAACCCAGGGCTCTGCCCGACTGTTAGAAATATAGGTCTCTTCTCCAG	8280
Db	8221	tctccagctcagaaccocagggctctgccccagTcgtTgaaataaggTctctctctcccaag	8280
Qy	8281	AATCCCAGCGGCCCAATGGAACACCTCACGCTGGGTCTTAATTACAGTCTTTAAAGGCC	8340
Db	8281	aatcccaagcgccaatTgaaaacctcacgtggTcctaaTtaccagctctttaaagccc	8340
Qy	8341	AGCCCCTAGAAACCCAGTCTCTCTGGAACCCGCTACCTTAGAGCCAGACCAACGTTAC	8400
Db	8341	agccccTagaaccaccaagctcctcctcgggaaccgctccacttagagccagaccagcttac	8400

RESULT 2

RESULT
AAV57415

AAV57415
ID AAV57415 standard; cDNA; 2772 BP.

XX
XX

AC AAV57415;

XX

DT	11-JAN-1999 (first entry)
XX	
DE	Tumour suppressor gene MEN1 transcribed sequence (cDNA).

KW Familial multiple endocrine neoplasia type 1; FMEN1; MEN1;

KW menin; tumour suppressor gene; cancer; marker; diagnosis;
KW gene therapy; human; ss.
XX

YY Homo sapiens.

XX
XX

FH	Key	Location/Qualifiers
----	-----	---------------------

FT	CDS	111.1

FT 2211

FT: polyA_signal 2744..

/*tag= b

FT XX W09839439-A1.
XX PN
XX PD 11-SEP-1998.
XX PF 04-MAR-1998; 98WO-US04258.
XX PR 05-MAR-1997; 97US-0040269.
XX PA (USSH) US DEPT HEALTH & HUMAN SERVICES.
XX PI Agarwal SK, Burns AL, Chandrasekharappa SC, Collins FS;
PI Debelenko IV, Emmert-Buck MR, Guru SC, Liotta LA;
PI Lubensky IA, Manickam P, Marx SJ, Spiegel AM;
XX WPI: 1998-506360/43.
DR P-PSDB; AAW29749.
XX
PT Protein and gene associated with multiple endocrine neoplasia type 1
PT - useful in gene therapy and to diagnose sufferers of, and those
PT susceptible to, this condition by detecting protein absence or gene
PT mutation(s)
XX
PS Claim 7; Page 55-56; 75pp; English.
XX
CC This is the transcribed sequence (cDNA) of human MEN1, a novel
CC tumour suppressor gene which is associated with multiple endocrine
CC neoplasia type 1. MEN1 cDNA can be obtained e.g. from leukocyte
CC RNA using MEN1-specific probes and PCR primers. A genomic sequence
CC (see AAW57416) is also provided. MEN1 encodes a 67.5 kDa protein
CC (see AAW28749), termed menin, that exhibits no apparent similarities
CC to previously known proteins. The lack of a functional menin
CC polypeptide, either by absence of the protein, its alteration
CC and/or associated mutations in the MEN1 gene, have been identified
CC in individuals with familial multiple endocrine neoplasia type 1
CC (FMEN1) and suffering from multiple endocrine neoplasia type 1.
CC The identification of MEN1 provides a new window into the mechanism
CC of endocrine tumorigenesis, facilitates accurate early diagnosis of
CC MEN1 associated cancers, and provides preclinical identification of
CC individuals with the FMEN1 syndrome, i.e. cancer-free individuals
CC that are at high risk of acquiring MEN1 associated tumours. It
CC also provides a basis for gene therapy. The MEN1 nucleic acids may
CC also be used therapeutically to produce menin recombinantly and to
CC produce transgenic animals useful in research.
XX
SQ Sequence 2772 BP; 577 A; 909 C; 779 G; 507 T; 0 other;

Query Match 14.3%; Score 1312.2; DB 19; Length 2772;
Best Local Similarity 96.5%; Pred. No. 1.6e-234;
Matches 1341; Conservative 0; Mismatches 48; Indels 0; Gaps 0;

QY 7499 GGATGCCAGACGAGGTCTGGAGTTCAGCCACTGCGCGGCAACTTGCCTCACTT 7558
DB 1382 gaaaggcagtcacacgcctgtgtgcacgtggcgccaccttctgtgagtcct 1441
QY 7559 GCTCTCCCACTGGCCAGGTGCGGAGAGGTGCGCATAGTGAGCGGAGCGCGAGGC 7618
DB 1442 aggcctttttagggacaggtgcgcagaaaggtcgcatagtgagccgagagccgaggc 1501
QY 7619 GGCCGAGGCGGAGGAGCGGTGGGGGAGGAAGCCCGGAGCGCGCGGGGCCACG 7678
DB 1502 ggcggagccgagagccgtgggcgagaaagcccggaagccgagccgggcccacg 1561
QY 7679 GCGGGAGTCCAGCCAGAGAGGCCCCCGCCGAGAGAGCCAGAGCTGGACAGGGCCT 7738
DB 1562 gcgggagtccaagccagagagcccccgcgcgccaagaagagccagcactggacagggcct 1621
QY 7739 GGCACCGCGCAGGGTGCAGTGTACAGACCCCGGAGAGCCCTCCCTGGAGCTGCTGG 7798
DB 1622 gggcaccggccagggtgcagtgctcaggatcccccccggaagccctctggactgtcgtgg 1681

QY 7799 CACAGCCCGAGGCGCTGAAGGTGGCAGCAGCGCTCAGGTGCCAGCACCCCGACATCAC 7858
DB 1682 cacagcccgagggccctgaaggtggcagcagcgtcagggtgccagcaccgcagcacaacc 1741
QY 7859 ACCGCGGAGGGTCCAGTGTCTACTTTCCAGAGTGAGAGATGAAGGCGATGAAGAGCT 7918
DB 1742 accgcccggagggtccagtgctcactttccagagtgagaaagtgaaaggcagtgaggagct 1801
QY 7919 GCTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGCACAGTCGCA 7978
DB 1802 gctgggtggccaccaagatcaactcagcgcagccatcaagctgcaactcagcgcagtcgca 1861
QY 7979 AGTCAGATGAAGAAGCAGAAAGTGCCACCCCTAGTACTACACTCTGTCTTTCTCTCAA 8038
DB 1862 agtcagatgaagaagcagaaagtgtccaccctcctgactcacactctgtcttctctcaa 1921
QY 8039 GCGGACGCGCAAGGCGCTCTGAAGTCTTGGGACTTTCGGACCGCTTGTGGGGACCCAGGC 8098
DB 1922 gcggcagcgcgaaggcctctgaactactgggacttcggaccgcttgtggggaccagcgc 1981
QY 8099 TCGGCTTGTAGTCCCGCAACTCTGAGCCCAATGTCGCCCGCCAGCCCAAGGGGACAGGCC 8158
DB 1982 tcgcctctagtcggcccaactctgagcccatgttctgccccagcccaaggggacagggcc 2041
QY 8159 TCACCTCTACCCAAACCCCTAGGTTCGCGGTCCCGAGTACAGTCTGTATCAAAACCCAGAT 8218
DB 2042 tcactctaccacaacccctaggttcccggtcccgagtagactgtatcaaacccagat 2101
QY 8219 TTTCTCCAGCTCAGAACCCAGGCGCTCTGCCCGAGTCTGTAGATAATAGTGCTCTTCTCC 8278
DB 2102 ttctccagctcagaaacccagggctctgccccagtgcttagaataaggtctctctctccc 2161
QY 8279 AGAATCCAGCGCGGCAAGTGGAACTCAGCTGCGGTGGTCTTAATTACAGCTTTTAAAGGC 8338
DB 2162 agaatccccgcggcccaatggaaacctcagcgtgggtctcaattaccagtccttaaggcc 2221
QY 8339 CCAGCCCTTAGAAACCCAAAGCTCCTCTCGGAACCGCTCACCTAGAGCCAGACCAAGCTT 8398
DB 2222 ccagcccttagaaacccagctcctcctcgaaaccgctcactcactagagcagacacagtt 2281
QY 8399 ACTCAGGCTCTCCCGAGCTGTAGAGCTGAGGTTTCACCCCTTAACCCAAAGGAGGACACA 8458
DB 2282 actcagggctcctcccagctgttagagctgaggtgttccaccttaacccagggagcaca 2341
QY 8459 GGTCCACACTCCAGCGCGGAGCCCTAGGACCACTCAGGCCCTTAGGATATATTTCGCA 8518
DB 2342 ggtcccaactccagcccggggagccttaggaccactcagcccttaggagtagatttccgca 2401
QY 8519 CTTAGAAATTCATATCTTTCGGAATCCAAAGTCCCTGCCCAATAACTTCACTCTCTGCT 8578
DB 2402 cttcagaattccatatctgtcggaatccaagctccctgcccccaataacttcagtcctgct 2461
QY 8579 TCCAGAAATTTGAAATTCCTAGTTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 8638
DB 2462 tcagaatttggaatctcctagtttctcctctcctcctcctcctcctcctcctcctcctcct 2521
QY 8639 CCGCCCGCAGCTATGAGCATCTGAGCCCGCCGCTCTCTCTGACCAAACTGCGCCCGGA 8698
DB 2522 cgcgccccagcctatgagcactcctgagccccgcctctcctgacgcgaactggccccgga 2581
QY 8699 TCAGAGCAGGACCTCCCTTCCGACCTCTGGGAACCTTCCAGAGGTCCAGCCCATCTCGG 8758
DB 2582 tcagagcaggaaacctcctccgacctctggaaacctcccagagggtccagccccatctcgg 2641
QY 8759 AGCATCCCGAGGAATCTGCAGAGGGGTAGGAGTGGGTGACAAAGAGCTGATCTCTTC 8818
DB 2642 agcatccccgaggaactctgcagaggggttaggagtggtgacaagagcctgatctcttc 2701
QY 8819 CTCTTTGTACATAGATTTATTTTTCAGTTCCAGAAAGATGAATACATTTTGTAAAAA 8878
DB 2702 ctgtttgtacatagattttatttttcagttccccagaagaagatgaatacatatttgttataaaa 2761
QY 8879 AAATATAAA 8887

PN WO9854324-A1.
XX 03-DEC-1998.
XX 29-MAY-1998; 98WO-US10957.
XX 29-MAY-1997; 97US-0865337.
XX (INCY-) INCYTE PHARM INC.
XX Au-Young J, Covitz PA, Murry LE, Tang YT;
XX WPI; 1999-059839/05.
XX New isolated multiple endocrine neoplasia type 1 protein - used to
PT develop products for the diagnosis, treatment and prevention of
PT cancers and multiple endocrine neoplasia.
XX Disclosure; Page 49; 67pp; English.
XX The present sequence represents Incyte clone 1313372 which is used to
CC make up the consensus human multiple endocrine neoplasia type 1 protein
CC (MEND-1) (see AAV80659). MEND-1 plays a role in multiple endocrine
CC neoplasia when one or both normal genetic copies of MEND-1 are mutated
CC and no longer able to suppress tumourigenesis. MEND-1 can be used for
CC treating or preventing cancers and multiple endocrine neoplasia. MEND-1
CC polynucleotides can also be used for gene therapy. Products from the
CC present invention can also be used for detection, diagnosis and drug
CC screening.
XX Sequence 812 BP; 188 A; 270 C; 209 G; 131 T; 14 other;
SQ

Query Match 8.1%; Score 746.4; DB 20; Length 812;
Best Local Similarity 97.2%; Pred. No. 1.2e-129;
Matches 772; Conservative 4; Mismatches 16; Indels 2; Gaps 2;

QY 7687 CCAAGCCAGAGAGCCGCCCCGCCCAAGAGCCAGCCTGGACAAGGGCTGGGCACCC 7746
DB 1 ccaagcagagagcccccgcgcacaaagaagccagcactggacaaggcctggcaccg 60
QY 7747 GCCAGGTCGAGTGTGAGGACCCCGGGAAGCCTCTGCGGACTGTGCTGGGCACGCC 7806
DB 61 gccaggcgcagtgctcaggaccgccgcggaagcctcctgggaactgtcgtggcacagcc 120
QY 7807 GAGGCCCTGAAGTGGCGAGCAGCGCTCAGTGCAGCACCGCAGCATCACACCGCGG 7866
DB 121 gaggccctgaagtggcagcagcagcagcagcagcagcagcagcagcagcagcagcagc 180
QY 7867 AGGGTCCAGTGTCTCTCCAGAGTGAGAGATGAAGGCATGAAGGAGCTGCTGGTGG 7926
DB 181 agggctcagtgctcactttccagagtgagaagatgaaggscatgaaggagctgctgtgg 240
QY 7927 CCACCAAGATCAACTCGAGCGCCATCAAGCTGCAACTCAGCGGCACAGTCCCAAGTCCGA 7986
DB 241 ccaccaagatcaactcagcgcagcagcagcagcagcagcagcagcagcagcagcagcagc 300
QY 7987 TGAAGAAGCAGAAGTGTCCACCCCTAGTGTACTACTGTCTTCTTCTCAAGCGGCAGC 8046
DB 301 tgaagaagcagaagtgctccacccttagtgactacactctgtcttctcctcaagcgcag 360
QY 8047 GCAAAGGCTCTGAACACTACTGCGGACTTCGACCGCTGTGGGACCCAGGCTCCG-CCT 8105
DB 361 caaaaggcctctgaactactcgggacttcggaccgctgttgggaccagcctccgacct 420
QY 8106 TAGTCCCGCAACTCTGAGCCCATGTCTGCCCCAGCCCAAGGGGACAGGCGCTCAGCTC 8165
DB 421 tagtcccccaactctgagcccatgtctgtccccagccccaaaggagagcctcactc 480
QY 8166 TACCCAAAACCTTAGGTCCCGGTCGCGAGTACAGTGTGTATCAAAACCCAGATTCTCC 8225
DB 481 tacccaaaccttaggttcccggtcccgagtacagtgtgtatcaaaacccagatttctcc 540

QY 8226 AGCTCAGAACCCAGGGCTCTGCCCCAGCTCGTTAGATATAGGTCCTTCTCCCAAGATCC 8285
DB 541 agctcagaacccagggctctgccccagctgtagaataataggctctcttccccagaatcc 600
QY 8286 CAGCCGGCCAATGAAACCTCAGCTGGTCTTAATTACAGTCTTTAAAGGCCACGCC 8345
DB 601 cagccggcmaatgaaacacctcagctgggtcctaatctaccagtcctttaaaggccagccc 660
QY 8346 CTAGAAAACCCAAAGTCTCTCTCGG-AACCGCTCACCTAGAGCCAGACCAAGTTACTCAG 8404
DB 661 ctagaaaacccaaagctctctcggnaacctgtcaccagagccagaccagcttctaytcaag 720
QY 8405 GGCTCTCTCCAGCTGTGAGGAGCTGAGTTTCACCCCTTAACCCCAAGGAGCAGAGTCC 8464
DB 721 gggctctnccagctgttaggagctgaggttttnacccttnacccaaggaggaanggtccc 780
QY 8465 ACCTCCAGCCCGGG 8478
DB 781 ancttcagcccg 794

RESULT 5
AAV80660
ID AAV80660 standard; cDNA; 538 BP.
XX
XX AAV80660;
XX AC
XX 15-MAR-1999 (first entry)
XX
XX Human multiple endocrine receptor neoplasia type 1 clone 1596949.
DE Human; multiple endocrine neoplasia type 1 protein; MEND-1;
XX tumorigenesis; cancer; multiple endocrine neoplasia; gene therapy;
KW detection; diagnosis; drug screening; ss.
XX
XX Homo sapiens.
XX
XX WO9854324-A1.
XX
XX 03-DEC-1998.
XX
XX 29-MAY-1998; 98WO-US10957.
XX
XX 29-MAY-1997; 97US-0865337.
XX
XX (INCY-) INCYTE PHARM INC.
XX
XX Au-Young J, Covitz PA, Murry LE, Tang YT;
XX WPI; 1999-059839/05.
XX
XX New isolated multiple endocrine neoplasia type 1 protein - used to
PT develop products for the diagnosis, treatment and prevention of
PT cancers and multiple endocrine neoplasia.
XX
XX Disclosure; Page 48; 67pp; English.
XX
XX The present sequence represents Incyte clone 1596949 which is used to
CC make up the consensus human multiple endocrine neoplasia type 1 protein
CC (MEND-1) (see AAV80659). MEND-1 plays a role in multiple endocrine
CC neoplasia when one or both normal genetic copies of MEND-1 are mutated
CC and no longer able to suppress tumourigenesis. MEND-1 can be used for
CC treating or preventing cancers and multiple endocrine neoplasia. MEND-1
CC polynucleotides can also be used for gene therapy. Products from the
CC present invention can also be used for detection, diagnosis and drug
CC screening.
XX
XX Sequence 538 BP; 77 A; 197 C; 150 G; 110 T; 4 other;
SQ

Query Match 5.1%; Score 471.4; DB 20; Length 538;
Best Local Similarity 98.3%; Pred. No. 1.1e-78;
Matches 475; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

RESULT 7
AAV27941/c
ID AAV27941 standard; cDNA; 14796 BP.
XX AC AAV27941;
XX DT 25-SEP-1998 (first entry)
XX DE Survivin gene.
XX KW survivin; apoptosis; cellular apoptosis; transplantation; ss;
KW motor neuron degenerative disease; HIV infection; immunosuppression;
KW gastrointestinal perturbations; cardiovascular disorder.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FT CDS 2811..12044
FT FT /*tag= a
FT FT /product= Survivin
FT FT 2811..2921
FT FT /*tag= b
FT FT /number= 1
FT FT 2922..3173
FT FT /*tag= c
FT FT /number= 1
FT FT 3174..3283
FT FT /*tag= d
FT FT /number= 2
FT FT 3284..5157
FT FT /*tag= e
FT FT /number= 2
FT FT 5158..5275
FT FT /*tag= f
FT FT /number= 3
FT FT 5276..11954
FT FT /*tag= g
FT FT /number= 3
FT FT 11955..12044
FT FT /*tag= h
FT FT /number= 4
XX PN WO9822589-A2.
XX DT 28-MAY-1998.
XX PD 20-NOV-1997; 97WO-US21880.
XX PF 20-NOV-1997; 97US-0975080.
XX PR 20-NOV-1996; 96US-0031435.
XX PA (UYVA) UNIV YALE.
XX PI Altieri DC;
XX DR WPI; 1998-312475/27.
XX DR P-PSDB; W614359.
XX PT Modulating apoptosis by controlling the Survivin gene - useful for
XX PT treating transplant rejection, degenerative disorders and tumours
XX PS Disclosure; Fig 10; 108pp; English.
XX CC The survivin gene can be used to control apoptosis through modification
XX CC of the gene. Survivin peptides can be used to inhibit cellular
XX CC apoptosis, e.g. for enhancing the viability of organs and tissues prior
XX CC to their transplantation, for preserving the growth of cells in culture
XX CC or for treating conditions involving abnormal apoptosis,
XX CC e.g. degenerative diseases such as motor neuron degenerative diseases,
XX CC HIV infection, dermatological effects of ageing, disorders and diseases
XX CC such as immunosuppression, gastrointestinal perturbations, cardiovascular
XX CC disorders, apoptosis related to reperfusion damage, rejection of tissue
XX CC transplantation and Alzheimer's disease. Agents which block Survivin

CC activity can be used to treat e.g. tumours.
XX Sequence 14796 BP; 3482 A; 3488 C; 3888 G; 3938 T; 0 other.
SQ
Query Match 3.8%; Score 348; DB 19; Length 14796;
Best Local Similarity 66.8%; Pred. No. 1.8e-55;
Matches 624; Conservative 0; Mismatches 285; Indels 25; Gaps 8;
QY 382 GCACGACAGTGGCTCACACCTGTATCCACACCTTTTGGAGGCGGAGGTAGGAGAT 441
DB 6289 GCGGGTGGTGGCTCACACCTGTATCCACACCTTTGGAGGCGGAGGTAGGAGAT 6230
QY 442 CACTTGAGGTGAGAGTTCGAGACCCCTGACCAACATGGTGACACCCCTGTCTCTACT 501
DB 6229 CAC--GAGGTGAGAGATCAAGACCCCTCTGGCTAACACGGTGAAACCG-GTCTCTACT 6173
QY 502 AAGTAAATACAAAATTAGCCAAAGTGTGGTGGCAGCACCTGTAAATCCGGGTACTTTGG 561
DB 6172 AA-AAATACAAAATTTAGCCGGTGTGGTGGGCGGCTGTAGTCCAGCTACTCGG 6114
QY 562 GAAGCTGAGGCAGAGAATCACTTGAACCTGGGAGGCGAGAGTTGCAGTGAGCCGAGATC 621
DB 6113 GAGGCTGAGGCAGAGAATGGCTTGAACCTGGGCGGCGAGCTTGCAGTGAGCCGAGATC 6054
QY 622 ACCCCACTGCACCTCCAGCCTGAGTGACAGAGCGAGACTCTGTCTCAAAACAA----- 673
DB 6053 CCGTCGCTGCACCTCTAGCCTGGGCGACAGAGCAAGACTCCATCTCAATAAATAATAA 5994
QY 674 ---ATAAACAAATAACTACTCTTTGGCCGGTAAAGTGGTTCACGCTGTAAATTTAGCA 730
DB 5993 TTTAAAAATAAAATAACTTCTCAGGCCAGGTGTGGTGGCTCACACCTATAATCTCAGCA 5934
QY 731 CTTTGGAGGCTGAGGCGGCGACATCACTTGAAGTTAGGGTTTCGAGACCAAGTCTGCCCA 790
DB 5933 CTTTGGAGGCTGAGGCGGCGTGGATCAGCTGAGGTTCAGAGTTTCAAAACAGCGCTGCCCA 5874
QY 791 ACATGGTAAACCCCATCTCTACTTAAATAACAAAAGTTTCTGGGTGTGGTGGCGGAC 850
DB 5873 ACATGGTAAACCTCGTCTCTACCAAAATAACAAA--TTACCTGGTGTGGTGGCACAT 5815
QY 851 GCCTATAATCCAGCTACTTTGGGACTTTTAAAGACGGAATCTCACTCTGTGTGGCCAG 910
DB 5814 GCCTGTAAATCTCGCTACTTTGGGAGGCT-----GAAGCAGGCAAAATCACTTGAGCCCA 5762
QY 911 GCTGGAGTGCAGTGGCAAGATTCTGGCTCACTGAAGCTCCGCTCCAGGTTCAAGGGG 970
DB 5761 GGAGCGGAGGTTTGCAGTGTATCGAGATCAGGCCACCTACTCCAGCTCGGTAAACAGAG 5702
QY 971 ATTCCCGCGGCTCAGCCTCCCAAGTAGCTGGGAATCCCTGTCTCTGCAAAAAAATAA 1030
DB 5701 TGAACCTGTCTCAGAACAAAAAAGCCAGCGGTGGTGGTTCAGCCCTCTAATC 5642
QY 1031 AAAAAAATAAATAAT 1090
DB 5641 CCAGCACTTTGGGAGCCCAAGGTGGCGAATCAGAGGTTCAGAGTTCAAGACCAAGCTG 5582
QY 1091 TTATATGTAT 1150
DB 5581 GGCATATGTGTGAAACCCCATCTCTACTTAAATAATACA-AAATATAGCTGGGTGGTGGC 5523
QY 1151 GCACGCTGTGATCACAGCTACTCGGAGGCTGAGGCAGCAGAGATCGTTGAGCCCGTGA 1210
DB 5522 ACTTCCTGTATATCCAGTTACTTGGAGGCTGAGGCAGGAGAGTAATCACTTGAACCCGAGA 5463
QY 1211 AGTCGAGGTGAGTGGAGCCAGATCGAGCCACTGCTATTCAGCCTGGGCGA-AAGAGAA 1269
DB 5462 GGCAGAGGTGAGTGGAGCCAGATTCACCAATTGCACTCCAGCCTCGGCAACAAGAGTG 5403
QY 1270 AGACCTGTCTCAAAACAACAACAACAAGCTAC 1303
DB 5402 AAACCTATCTCAGAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 5369

OS Homo sapiens.
PN WO200116311-A1.
XX
PD 08-MAR-2001.
XX
PF 31-AUG-2000; 2000WO-JF05930.
XX
PR 31-AUG-1999; 99JP-0245962.
PR 09-MAY-2000; 2000JP-0136266.
XX
PA (HISM) HISAMITSU PHARM CO LTD.
PA (CHIB-) CHIBA PREFECTURE.
XX
XX Nakagawara A;
XX WPI; 2001-226686/23.
XX
XX Human lp36 homozygosity deletion domain from the 36-position of first
PT chromosome short arm in human neuroblastoma cell lines, applicable e.g.
PT in gene diagnosis of tumors as well as in developing anti-cancer drugs
PT
XX
PS Example 8; Page 67-88; 226pp; Japanese.
XX
CC The present invention describes a homozygosity deletion domain
CC co-existing in the 36-position of the first chromosome short arm (lp36)
CC in human neuroblastoma. Also described are base sequences from the lp36
CC position of human neuroblastoma cell lines (NB-1 and MASS-NB-SCH-1),
CC which are tumour suppressor genes in human neuroblastoma. The genes are
CC tumour suppressor genes, base sequence data of which are applicable as
CC tumour markers and reagents in studying mechanism of tumour body
CC formation, and gene diagnosis of tumours as well as in developing
CC anti-cancer drugs. AAF9787 to AAF97829 represent PCR primers used in
CC the exemplification of the present invention, and AAF97830 to AAF97874
CC represent sequences given in the exemplification of the present
XX invention.
XX
SQ Sequence 51474 BP; 14008 A; 11363 C; 11657 G; 14446 T; 0 other;

Query Match 3.7%; Score 342.8; DB 22; Length 51474;
Best Local Similarity 78.8%; Pred. No. 2.1e-54;
Matches 490; Conservative 0; Mismatches 112; Indels 20; Gaps 6;

QY 3121 TTTTGTGAGACAGGCTCTCACCTCTCCGCCAGGCTGGAGTGCAGTGTGTG 3180
DB 15291 TTTTGTGAGACAGAGTCTCGCTCTCGGCCAGGCTGGAGTGCAGTGGCGG 15232

QY 3181 ATCATGGCTCAGTGCAGGCTCAGTCTCCCTGGGCTCAGAGATCCTCCAACTCAGCCTCC 3240
DB 15231 ATCTGGCTCAGTGCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAG 15172

QY 3241 TGAGTAGCTCAGAGTACAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAG 3299
DB 15171 CGAGTAGCTGGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAG 15112

QY 3300 GTAGAGATGGTGTCTCACTATGTGGCCAGACTGGTCTTCAACTCTTGGGCTCAAGTGAT 3359
DB 15111 GTAGAGATGGGTTTACCAGATTTAGCCAGATGGTCTCAGATCTCTGACCTC-TGTGAT 15054

QY 3360 CTGCGGCTCAGTCTCC-AAAATGCTGGGATTTACAGGCTGAGCCACCCAGCTGGTGG 3418
DB 15053 CGCGGCTCAGCTCCCAAAATGTTAGGATTTACAGGCTGAGCCACTCGGCCAGCTC 14994

QY 3419 CCTATGAAATTTTTTTTTTTTTCAGACGGGCTCTCAGTCTGCGCCAGGCTGGAGTG 3478
DB 14993 ATTTT-----TTTTTTTTTTTCTAGATGAGGCTCGCTGTGTGTTGCCAGGCTGGAGTG 14939

QY 3479 CAGTGGTGAATCTCGGCTCAGTGCAGGCTCTGCTCTCTCTCATGCCATTTCTCTCGC 3538
DB 14938 CAGTGGCACAATCTCGGCTCAGTGAACCTCCGCTCCCAAGATTCAGCGATT----- 14884

QY 3539 CTCTGCTCAGCCTCCTGAGTAGTGGACTACAGGAGCTGCCACCATGCTGGCTAA 3598
DB 14885 CTCTGCTCAGCCTCCTGAGTAGTGGACTACAGGAGCTGCCACCATGCTGGCTAA 14826

QY 3599 TTTTGTGAGATTTTGTAGTAGAGAGAGGTTTCACCATGTTAGCCAGGATGGTCTCGAT 3658
DB 14825 ----CTTTGTATTTTGTAGTAGAGATGGGTTTTCACCATGTTGGCCAGGATGGTCTTGTAT 14770

QY 3659 CTCTGACCTCGTGTCCGCCCTTGGCTCCCAAGTGTGGGATTACAGGCTGAG 3718
DB 14769 CTTTGTACCTCGTGTGTTGCCACCTTGGCTCCCAAGTGTGGGATTACAGGCTGAG 14710

QY 3719 CCACCGCACCTGCTCAAAATG 3740
DB 14709 CCATTGCACCCAGCTTACAGTG 14688

RESULT 15
AAV27941
ID AAV27941 standard; cDNA; 14796 BP.
XX
AC AAV27941;
XX
DT 25-SEP-1998 (first entry)
XX
DE Survivin gene.
XX
KW survivin; apoptosis; cellular apoptosis; transplantation; ss;
KW motor neuron degenerative disease; HIV infection; immunosuppression;
KW gastrointestinal perturbations; cardiovascular disorder.
XX
OS Homo sapiens.
XX
FH Key
FT CDS 2811..12044
FT /tag= a
FT /product= Survivin
FT 2811..2921
FT /tag= b
FT /number= 1
FT /tag= c
FT /number= 1
FT /tag= d
FT /number= 2
FT /tag= e
FT /number= 2
FT /tag= f
FT /number= 3
FT /tag= g
FT /number= 3
FT /tag= h
FT /number= 4

PN WO9822589-A2.
XX
PD 28-MAY-1998.
XX
PF 20-NOV-1997; 97WO-US21880.
XX
PR 20-NOV-1997; 97US-0975080.
PR 20-NOV-1996; 96US-0031435.
XX
PA (UYUA) UNIV YALE.
XX
PI Altieri DC;
XX
DR WPI; 1998-312475/27.

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 14, 2001, 01:32:59 ; Search time 15359.2 Seconds
(without alignments)
9244.854 Million cell updates/sec

Title: US-09-380-337-3
Perfect score: 9180
Sequence: 1 CTGGTCTTGAACCTCTGGCC.....AGCCACGGCGCGCGCCGCC 9180

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues 2688314
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database :
- GenEmbl:*
 - 1: gb_ba1:*
 - 2: gb_ba2:*
 - 3: gb_ba3:*
 - 4: gb_in1:*
 - 5: gb_in2:*
 - 6: gb_in3:*
 - 7: gb_om:*
 - 8: gb_ov:*
 - 9: gb_pat1:*
 - 10: gb_pat2:*
 - 11: gb_ph:*
 - 12: gb_pl1:*
 - 13: gb_pl2:*
 - 14: gb_pl3:*
 - 15: gb_pl4:*
 - 16: em_ba1:*
 - 17: em_ba2:*
 - 18: em_fun:*
 - 19: em_htgo_hum:*
 - 20: em_htgo_inv:*
 - 21: em_htgo_rod:*
 - 22: em_htg_hum1:*
 - 23: em_htg_hum2:*
 - 24: em_htg_hum3:*
 - 25: em_htg_hum4:*
 - 26: em_htg_hum5:*
 - 27: em_htg_hum6:*
 - 28: em_htg_hum7:*
 - 29: em_htg_hum8:*
 - 30: em_htg_inv1:*
 - 31: em_htg_inv2:*
 - 32: em_htg_inv3:*
 - 33: em_htg_rod:*
 - 34: em_hum1:*
 - 35: em_hum2:*
 - 36: em_hum3:*
 - 37: em_hum4:*
 - 38: em_hum5:*
 - 39: em_hum6:*
 - 40: em_hum7:*
 - 41: em_in:*
 - 42: em_om:*
 - 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vi:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vil:*
- 59: gb_vl2:*
- 60: gb_vl3:*
- 61: gb_vl4:*
- 62: gb_vl5:*
- 63: gb_vl6:*
- 64: gb_vl7:*
- 65: gb_vl8:*
- 66: gb_vl9:*
- 67: gb_vl10:*
- 68: gb_vl11:*
- 69: gb_vl12:*
- 70: gb_vl13:*
- 71: gb_vl14:*
- 72: gb_vl15:*
- 73: gb_vl16:*
- 74: gb_vl17:*
- 75: gb_vl18:*
- 76: gb_vl19:*
- 77: gb_vl20:*
- 78: gb_vl21:*
- 79: gb_vl22:*
- 80: gb_vl23:*
- 81: gb_vl24:*
- 82: gb_vl25:*
- 83: gb_vl26:*
- 84: gb_vl27:*
- 85: gb_vl28:*
- 86: gb_vl29:*
- 87: gb_vl30:*
- 88: gb_vl31:*
- 89: gb_vl32:*
- 90: gb_vl33:*
- 91: gb_vl34:*
- 92: gb_vl35:*
- 93: gb_vl36:*
- 94: gb_vl37:*
- 95: gb_vl38:*
- 96: gb_vl39:*
- 97: gb_vl40:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	9179	100.0	9180	97	HSU93237
C	2	9042	98.5	203300	85 AC000134 Human menin
3	8004.8	87.2	109290	84	HS8381.14 Y12335 Homo sapien
C	4	5844	63.7	171980	82 AP000928 Homo sapi
C	5	5072.2	55.3	73431	82 AP000575 Homo sapi
C	6	4368.4	47.6	166269	82 AP001187 Homo sapi
7	4118.2	44.9	157454	82	AP001462 Homo sapi
8	3386.2	36.9	178100	82	AP001558 Homo sapi

9	2071.4	22.6	73431	82	AP000575	AP000575 Homo sapi
10	1982.2	21.6	166269	82	AP001187	AP001187 Homo sapi
11	1312.2	14.3	2772	97	HS093236	U93236 Human menin
12	1253.2	13.7	1691	91	BC002664	BC002664 Homo sapi
13	1251.6	13.6	2633	91	BC002544	BC002544 Homo sapi
14	1242.4	13.5	1837	97	HS093236	Y12338 H.sapiens m
15	1075.8	11.7	9286	94	AF109390	AF109390 Mus muscu
16	1075.8	11.7	132365	94	AC006956	AC006956 Mus muscu
17	1058.2	11.5	6779	94	AF093756	AF093756 Mus muscu
18	1036.6	11.3	6736	94	AF024513	AF024513 Mus muscu
19	942.4	10.3	171980	82	AP000928	AP000928 Homo sapi
20	942.	10.3	178100	82	AP001558	AP001558 Homo sapi
21	782.6	8.5	157454	82	AP001462	AP001462 Homo sapi
22	776.6	8.5	938	93	HS093236	AJ297487 Homo sapi
23	749.4	8.2	2196	9	AR081879	AR081879 Sequence
24	746.4	8.1	812	9	AR081882	AR081882 Sequence
25	565.2	6.2	71297	78	AC091082	AC091082 Homo sapi
26	545.8	5.9	71297	78	AC091082	AC091082 Homo sapi
27	543.6	5.9	52358	77	AC090316	AC090316 Homo sapi
28	536.6	5.8	3067	94	AB023400	AB023400 Rattus no
29	525	5.7	2984	94	AF109389	AF109389 Mus muscu
30	513.4	5.6	2947	94	AF130368	AF130368 Mus muscu
31	483.4	5.3	944	93	HS093236	AJ297488 Homo sapi
32	477.8	5.2	2261	94	AF130369	AF130369 Rattus no
33	471.4	5.1	538	9	AR081880	AR081880 Sequence
34	466	5.1	501	93	HS093236	AJ297486 Homo sapi
35	465.6	5.1	816	93	HS093236	AJ297489 Homo sapi
36	465.4	5.1	513	93	HS093236	AJ297485 Homo sapi
37	456.2	5.0	2623	94	AF016398	AF016398 Mus muscu
38	428.8	4.7	2577	94	AF072755	AF072755 Mus muscu
39	408.8	4.5	1964	94	AB023401	AB023401 Mus muscu
40	396.6	4.3	1906	94	AF130370	AF130370 Rattus no
41	395.8	4.3	181683	66	AC021618	AC021618 Homo sapi
42	395.4	4.3	49868	78	AC090368	AC090368 Homo sapi
43	393.4	4.3	207091	77	AC087289	AC087289 Homo sapi
44	393	4.3	154978	75	AC073952	AC073952 Homo sapi
45	393	4.3	200491	86	AC007249	AC007249 Homo sapi
ALIGNMENTS						
RESULT	1	HSU93237	9180 bp	DNA	PRI	19-APR-1997
LOCUS	HSU93237	Human menin (MEN1) gene, complete cds.				
DEFINITION	U93237	U93237.1 GI:1945388				
ACCESSION	U93237.1	GI:1945388				
VERSION	human.					
KEYWORDS	human.					
SOURCE	human.					
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1	(bases 1 to 9180)				
AUTHORS	Chandrasekharappa,S.C., Guru,S.C., Manickam,P., Olufemi,S.-E., Collins,F.S., Emmert-Buck,M.R., Debeelenko,L.V., Zhuang,Z., Lubensky,I.A., Liotta,L.A., Crabtree,J.S., Wang,Y., Roe,B.A., Weisenmann,J., Boguski,M.S., Agarwal,S.K., Kester,M., Kim,Y.S., Heppner,C., Dong,Q., Spiegel,A.M., Burns,A.L. and Marx,S.J.					
TITLE	Positional cloning of the gene for multiple endocrine neoplasia-type 1					
JOURNAL	Science	276 (5311), 404-407 (1997)				
MEDLINE	97258940					
REFERENCE	2	(bases 1 to 9180)				
AUTHORS	Collins,F.S.					
TITLE	Direct Submission					
JOURNAL	Submitted (13-MAR-1997)	National Human Genome Research Institute, Bldg 38A, Room 605, National Institutes of Health, Bethesda, MD 20892, USA				
FEATURES	source	Location/Qualifiers				
	1..9180	/organism="Homo sapiens"				
		/db_xref="taxon:9606"				
repeat_region						
gene						
repeat_region						
repeat_region						
repeat_region						
exon						
mRNA						
mRNA						
exon						
exon						
CDS						
variation						

variation /standard_name="L22R"
/note="Lys22Arg"
/replace="g"
2534..2537
/gene="MEN1"
/standard_name="357del4"
/note="4 bp deletion"
/replace=""
2593
/gene="MEN1"
/standard_name="416delC"
/note="1 bp deletion"
/replace=""
2641..2643
/gene="MEN1"
/standard_name="K119del"
/note="3 bp deletion"
/replace=""
2689
/gene="MEN1"
/standard_name="512delC"
/note="1 bp deletion"
/replace=""
complement(3117..3725)
/rpt_type="dispersed"
/rpt_family="ALU"
4297..4505
/gene="MEN1"
/number=3
4445
/gene="MEN1"
/standard_name="W198X"
/note="Trp198stop"
/replace="a"
4476..4479
/gene="MEN1"
/standard_name="735del4"
/note="4 bp deletion"
/replace=""
4716..4844
/gene="MEN1"
/number=4
5177..5217
/gene="MEN1"
/number=5
5298..5385
/gene="MEN1"
/number=6
6025..6161
/gene="MEN1"
/number=7
6134
/gene="MEN1"
/standard_name="1132delG"
/note="1 bp deletion"
/replace=""
6623..6758
/gene="MEN1"
/number=8
6660..6662
/gene="MEN1"
/standard_name="E363del"
/note="3 bp deletion"
/replace=""
7196..7360
/gene="MEN1"
/number=9
7316
/gene="MEN1"
/standard_name="W436R"
/note="Trp436Arg"
/replace="c"
7317

exon /gene="MEN1"
/standard_name="W436X"
/note="Trp436stop"
/replace="a"
7578..7874
/gene="MEN1"
/number=10
7806
variation /gene="MEN1"
/standard_name="R527X"
/note="Arg527stop"
/replace="t"
BASE COUNT 1936 a 2604 c 2612 g 2027 t 1 others
ORIGIN

Query Match 100.0%; Score 9179; DB 97; Length 9180;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 9180; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CTGGTCTTGAACCTCTGGCCTCAAGCAATCCTCCTCCTTCAGCTTCCCAAGTCTTGTAA 60
|||||
Db 1 CTGGTCTTGAACCTCTGGCCTCAAGCAATCCTCCTCCTTCAGCTTCCCAAGTCTTGTAA 60
|||||

QY 61 TTACAGGCATGAGCCTGGCATGAACCTTGACACTATTGAGATATACATACAGGATTTTG 120
|||||
Db 61 TTACAGGCATGAGCCTGGCATGAACCTTGACACTATTGAGATATACATACAGGATTTTG 120
|||||

QY 121 TGAATGTCCTCAACTCTGTTTGGCAGATGTTTCTCATGATTAGAGGAGATTATAA 180
|||||
Db 121 TGAATGTCCTCAACTCTGTTTGGCAGATGTTTCTCATGATTAGAGGAGATTATAA 180
|||||

QY 181 ATTTTGGAGAAATCCAGAGAGGTGAAGAGTGAAGTAGGGCAGAAATTTAATCTGTTT 240
|||||
Db 181 ATTTTGGAGAAATCCAGAGAGGTGAAGAGTGAAGTAGGGCAGAAATTTAATCTGTTT 240
|||||

QY 241 ATTTACTGTATATACCGAGTGTCTGGAACCTTGCCCATGGTAAGTACCAAAATCTGTT 300
|||||
Db 241 ATTTACTGTATATACCGAGTGTCTGGAACCTTGCCCATGGTAAGTACCAAAATCTGTT 300
|||||

QY 301 TTTTGTGAATGAATAAGCAATAAATAGTGACCGTGGAAATTTAGTATTATTTCAAAGT 360
|||||
Db 301 TTTTGTGAATGAATAAGCAATAAATAGTGACCGTGGAAATTTAGTATTATTTCAAAGT 360
|||||

QY 361 TTCAAGCGTTGTTGATACAGGCCAGGCACAGTGGCTCACACCTGTAATCCCAGCACTTT 420
|||||
Db 361 TTCAAGCGTTGTTGATACAGGCCAGGCACAGTGGCTCACACCTGTAATCCCAGCACTTT 420
|||||

QY 421 TGGAGGCCGAGGTAGGAGGATCACTTGGGTGAGGAGTTTCGAGACGAGCCTGACCAACAT 480
|||||
Db 421 TGGAGGCCGAGGTAGGAGGATCACTTGGGTGAGGAGTTTCGAGACGAGCCTGACCAACAT 480
|||||

QY 481 GGTGACACCCCTGTCTCTACTAAGTAAAAATACAAAATTAGCCCAAGTGGTGGCAGGCA 540
|||||
Db 481 GGTGACACCCCTGTCTCTACTAAGTAAAAATACAAAATTAGCCCAAGTGGTGGCAGGCA 540
|||||

QY 541 CCTGTAATCCCGGCTACTTTGGGAAGCTGAGGCAGAGAAATCACTTGAACCTGGGAGGCAG 600
|||||
Db 541 CCTGTAATCCCGGCTACTTTGGGAAGCTGAGGCAGAGAAATCACTTGAACCTGGGAGGCAG 600
|||||

QY 601 AGGTTGCAGTGAGCCGAGATCACCCACCTGCACCTCCAGCCTGAGTGACACGAGCAGACTC 660
|||||
Db 601 AGGTTGCAGTGAGCCGAGATCACCCACCTGCACCTCCAGCCTGAGTGACACGAGCAGACTC 660
|||||

QY 661 TGTCTCAAAACAAATAAACAATAAATACTACTCTTTTGGCCGGTAAAGTGGTTCACGCCCTGT 720
|||||
Db 661 TGTCTCAAAACAAATAAACAATAAATACTACTCTTTTGGCCGGTAAAGTGGTTCACGCCCTGT 720
|||||

QY 721 AATTTTAGCATTCTGGGAGGCTGAGGGGGCAGATCACTTCAGGTTAGGGGTTTCGAGACC 780
|||||
Db 721 AATTTTAGCATTCTGGGAGGCTGAGGGGGCAGATCACTTCAGGTTAGGGGTTTCGAGACC 780
|||||

QY 781 AGTCTGGCCCAACATGGTGAAACCCCATCTCTACTTAAATACAAAAGTTTCTCGGCTGT 840
|||||

Db	5161	CTCCTTGGCTCTCTAGAACTGCTTGGCTGCTTATGACCTGGACAATCTGGAAAGGTC	5220
Qy	5221	AGTAGAGGAAGTGGCCAGGCTCGCCTGGTGAGGCCGGGGGCTGGGTGGCAGCCTGAA	5280
Db	5221	AGTAGAGGAAGTGGCCAGGCTCGCCTGGTGAGGCCGGGGGCTGGGTGGCAGCCTGAA	5280
Qy	5281	TTATGATCTTCTTCTAGTGTACCCATGGCCTTAGGCAACTGGCAGATCTAGAGAGCTG	5340
Db	5281	TTATGATCTTCTTCTAGTGTACCCATGGCCTTAGGCAACTGGCAGATCTAGAGAGCTG	5340
Qy	5341	GAGCCACCCCTGGCCGGCAGACCCACTCACCCCTTACCACAAGGTGGGGGCATCTAAG	5400
Db	5341	GAGCCACCCCTGGCCGGCAGACCCACTCACCCCTTACCACAAGGTGGGGGCATCTAAG	5400
Qy	5401	GAGGTGTCAGAGGAGACCCCTAACAGTGGCTGAGGCAGGGGCCCTCATCTGGGCAGATG	5460
Db	5401	GAGGTGTCAGAGGAGACCCCTAACAGTGGCTGAGGCAGGGGCCCTCATCTGGGCAGATG	5460
Qy	5461	AGAAGAGAACTTGTGTTTGGGGGTATCGCCCATTCAGTCTCACTTTGTGTCAACTGT	5520
Db	5461	AGAAGAGAACTTGTGTTTGGGGGTATCGCCCATTCAGTCTCACTTTGTGTCAACTGT	5520
Qy	5521	GTGCAGAACTCAGTTCAGTCAGGGCTGTCTGAGGGGTGTCCAGGGTCCCCAGCCTGGGAG	5580
Db	5521	GTGCAGAACTCAGTTCAGTCAGGGCTGTCTGAGGGGTGTCCAGGGTCCCCAGCCTGGGAG	5580
Qy	5581	TGCGAGGGCTGATTTGTTCCTCAGCCCTGCCTTTTCTGCCACTGCTTACTGTCTCTTC	5640
Db	5581	TGCGAGGGCTGATTTGTTCCTCAGCCCTGCCTTTTCTGCCACTGCTTACTGTCTCTTC	5640
Qy	5641	CTGGAGTATAACAGAGGTCAAAATGTGTAGAGCACTGAAGAGGGGTGTTCACCTTGGTG	5700
Db	5641	CTGGAGTATAACAGAGGTCAAAATGTGTAGAGCACTGAAGAGGGGTGTTCACCTTGGTG	5700
Qy	5701	GGTGTAGGTGGGAGAGGGCCATTGGGCTGGGCTTGAAGTCTTTGGTGTATGTGTAGAA	5760
Db	5701	GGTGTAGGTGGGAGAGGGCCATTGGGCTGGGCTTGAAGTCTTTGGTGTATGTGTAGAA	5760
Qy	5761	GAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGGGCAGGCCCCACCCCTGCAGTCTG	5820
Db	5761	GAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGGGCAGGCCCCACCCCTGCAGTCTG	5820
Qy	5821	CCCCAGGCTTCAGCCAGCAGTCTGTAGACCAGGGAGGACAGAGTGTAGAGGGCTGGC	5880
Db	5821	CCCCAGGCTTCAGCCAGCAGTCTGTAGACCAGGGAGGACAGAGTGTAGAGGGCTGGC	5880
Qy	5881	AGCGAGTGGAGTGGGAGTGGAGTGGAGAGACTCCCTGGGATCTTCCCTGTGGCCCCCTT	5940
Db	5881	AGCGAGTGGAGTGGGAGTGGAGTGGAGAGACTCCCTGGGATCTTCCCTGTGGCCCCCTT	5940
Qy	5941	CTGGGTGTGCCCTTGGTGGGCAATTTGTCCAGCAGGGCAGCTGGGGCTGCCTCCCTGAGG	6000
Db	5941	CTGGGTGTGCCCTTGGTGGGCAATTTGTCCAGCAGGGCAGCTGGGGCTGCCTCCCTGAGG	6000
Qy	6001	ATCCTCTGCCTCACCTCCATCCAGGCAATTGCCTCAGCCAGACCTACTATTCGGGATGAA	6060
Db	6001	ATCCTCTGCCTCACCTCCATCCAGGCAATTGCCTCAGCCAGACCTACTATTCGGGATGAA	6060
Qy	6061	CACATCTACCCCTACATGTACCTGGCTGGCTACCACTGTCCCAACCGCAATGTCCGGGAA	6120
Db	6061	CACATCTACCCCTACATGTACCTGGCTGGCTACCACTGTCCCAACCGCAATGTCCGGGAA	6120
Qy	6121	GCCTTCGAGGCTTGGCGGACAGGCCACTGTCTATCCAGGAGTGAAGATCCCCCTACTAG	6180
Db	6121	GCCTTCGAGGCTTGGCGGACAGGCCACTGTCTATCCAGGAGTGAAGATCCCCCTACTAG	6180
Qy	6181	GGCCTGCAGCCTGTCTTTCTTCCCTCCATCAGTTTCCACCACTCTGTCCAGGACTG	6240
Db	6181	GGCCTGCAGCCTGTCTTTCTTCCCTCCATCAGTTTCCACCACTCTGTCCAGGACTG	6240
Qy	6241	AGGCCTGGTCCCACGCCCATCCCTTTTCCATCCAGTTCCTTAGCGACAGGCCACCAT	6300

Db	6241	AGGCCTGGCTCCACAGCCCCCATCCCTTTCCATCCAGTCCCTTAGCGACAGAGCCACCAT	6300
Qy	6301	TACCAGAGGTAGGGACCCGTATTAAAGGTGTACATCTTTCCCTCCCTCCCCCTCTCTC	6360
Db	6301	TACCAGAGGTAGGGACCCGTATTAAAGGTGTACATCTTTCCCTCCCTCCCCCTCTCTC	6360
Qy	6361	CTAATTTTTTTTTTCTCAGAACAGTCTCAANTCTCCAATGTTTAAACCACCATCATCCAGC	6420
Db	6361	CTAATTTTTTTTTTCTCAGAACAGTCTCAAAATCTCCAATGTTTAAACCACCATCATCCAGC	6420
Qy	6421	AGTGGGACTTCCACCCCTCGCGCCCATGCCCCCTCCTCATTTCTGTTTCTTCCCTCTGGG	6480
Db	6421	AGTGGGACTTCCACCCCTCGCGCCCATGCCCCCTCCTCATTTCTGTTTCTTCCCTCTGGG	6480
Qy	6481	CTGACCCAGACAGATCATTTTTTGCACTGAGGACCCCACTACTCCCCCAGGCCCTTGGGG	6540
Db	6481	CTGACCCAGACAGATCATTTTTTGCACTGAGGACCCCACTACTCCCCCAGGCCCTTGGGG	6540
Qy	6541	CTCCATCCCCCGCAGTCCCTGGGGCTACCCCGATGCTGAGACCCCTTCAGACCCCTAC	6600
Db	6541	CTCCATCCCCCGCAGTCCCTGGGGCTACCCCGATGCTGAGACCCCTTCAGACCCCTAC	6600
Qy	6601	AGAGACCCACTGCTCTCAGACTACAATCTATGCGGGAGACGAGGAGATCTTACAAGG	6660
Db	6601	AGAGACCCACTGCTCTCAGACTACAATCTATGCGGGAGACGAGGAGATCTTACAAGG	6660
Qy	6661	AGTTCTTTGAAGTAGCCAAATGATCTATCCCAACTGCTGAAGGAGCGACCGAGCTTGC	6720
Db	6661	AGTTCTTTGAAGTAGCCAAATGATCTATCCCAACTGCTGAAGGAGCGACCGAGCTTGC	6720
Qy	6721	TGAGGCGGCGAGGAGCGGCCGGGGAGCAAGCCAGCTGAAAGGCTTGAGCTCCAGCC	6780
Db	6721	TGAGGCGGCGAGGAGCGGCCGGGGAGCAAGCCAGCTGAAAGGCTTGAGCTCCAGCC	6780
Qy	6781	TGTTCTCAGCCTCCCACCTGGACAGGGCTCCCTTCCACAGGGCCATGGGGCTGCATGTA	6840
Db	6781	TGTTCTCAGCCTCCCACCTGGACAGGGCTCCCTTCCACAGGGCCATGGGGCTGCATGTA	6840
Qy	6841	CGGGATTAGGGATGGCAGGAAAGTGGCCCTTGACAGACAGCTATCTTCCCTTTTGTCT	6900
Db	6841	CGGGATTAGGGATGGCAGGAAAGTGGCCCTTGACAGACAGCTATCTTCCCTTTTGTCT	6900
Qy	6901	ATACTCAGGTCTTGGGCCACGTTGGAGGGACTGAAAGTATTTTAGAGTTTCTACCC	6960
Db	6901	ATACTCAGGTCTTGGGCCACGTTGGAGGGACTGAAAGTATTTTAGAGTTTCTACCC	6960
Qy	6961	TGTGCTTCAGTTTCTATGGCCACACTCCCTCCTCAGCTGAGGGGTGGAGGTAGGATGG	7020
Db	6961	TGTGCTTCAGTTTCTATGGCCACACTCCCTCCTCAGCTGAGGGGTGGAGGTAGGATGG	7020
Qy	7021	TACGCTCGCTATGGATTGGCTTTTATAAAGGAAGAGGTTCTAAGAAATCTTCCCAACC	7080
Db	7021	TACGCTCGCTATGGATTGGCTTTTATAAAGGAAGAGGTTCTAAGAAATCTTCCCAACC	7080
Qy	7081	TATGCTTACCTTTCTTGGACCCAGGGGTCTTTGGCTAGGTGGGGGGCTGCGCTTGTGCC	7140
Db	7081	TATGCTTACCTTTCTTGGACCCAGGGGTCTTTGGCTAGGTGGGGGGCTGCGCTTGTGCC	7140
Qy	7141	TCTGCTAAGGGGTGAGTAAGAGACTCATCTGTGCCCTCCCTTCCCGCTGCTCCAGGGCAC	7200
Db	7141	TCTGCTAAGGGGTGAGTAAGAGACTCATCTGTGCCCTCCCTTCCCGCTGCTCCAGGGCAC	7200
Qy	7201	CCAGAGCCAGGTTTCGCGCTCCAGGACCCCTGAGTGTCTTCGCCCCACCTGCTGCGATTTCTA	7260
Db	7201	CCAGAGCCAGGTTTCGCGCTCCAGGACCCCTGAGTGTCTTCGCCCCACCTGCTGCGATTTCTA	7260
Qy	7261	CGACGGCATCTGAAATGGGAGGAGGCGAGTCCCACGCTGTGCTGCACGTGGGCTGGC	7320
Db	7261	CGACGGCATCTGAAATGGGAGGAGGCGAGTCCCACGCTGTGCTGCACGTGGGCTGGC	7320
Qy	7321	CACCTTTTGTGCAGTCCCTAGGCCGCTTTTATGGGACAGGTTGAGGGACAGCTCACAGA	7380
Db	7321	CACCTTTTGTGCAGTCCCTAGGCCGCTTTTATGGGACAGGTTGAGGGACAGCTCACAGA	7380

QY 7381 GGTCTGGCACTACAGTGGTGACAGACCCACGGGCTTGTGAGACTTTCTGGCCAGG 7440
|||||
Db 7381 GGTCTGGCACTACAGTGGTGACAGACCCACGGGCTTGTGAGACTTTCTGGCCAGG 7440
QY 7441 GGCAGCATCTGCCCATCCCTTCGTGGTGGCATGGGACTGAGACCCCTCGGGTGGGATGG 7500
|||||
Db 7441 GGCAGCATCTGCCCATCCCTTCGTGGTGGCATGGGACTGAGACCCCTCGGGTGGGATGG 7500
QY 7501 ATGGCCAGAGCAGGCTCTGGAGTTCCAGCCACTGSCCGCAACCTTGCCTCACTTGC 7560
|||||
Db 7501 ATGGCCAGAGCAGGCTCTGGAGTTCCAGCCACTGSCCGCAACCTTGCCTCACTTGC 7560
QY 7561 TCTCCCACTTGCCCAAGTGGCGCAAGAGTGCATAGTACCGCAGAGGCGGAGGGGG 7620
|||||
Db 7561 TCTCCCACTTGCCCAAGTGGCGCAAGAGTGCATAGTACCGCAGAGGCGGAGGGGG 7620
QY 7621 CGAGGCGAGAGCCGTGGGGCGAGAGAGCCCGGGAAGCGCGGCGGCGGCGCCACGGC 7680
|||||
Db 7621 CGAGGCGAGAGCCGTGGGGCGAGAGAGCCCGGGAAGCGCGGCGGCGGCGCCACGGC 7680
QY 7681 GGGAGTCCAAGCCAGAGAGCCCGCGCCGCCCAAGCCAGCACTGGACAGGCGCTGG 7740
|||||
Db 7681 GGGAGTCCAAGCCAGAGAGCCCGCGCCGCCCAAGCCAGCACTGGACAGGCGCTGG 7740
QY 7741 GCACCGCGCAGGCTGCTGAGGACCCCGCGGAAAGCCTCTGGGACTGTGCTGGCA 7800
|||||
Db 7741 GCACCGCGCAGGCTGCTGAGGACCCCGCGGAAAGCCTCTGGGACTGTGCTGGCA 7800
QY 7801 CAGCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCGCAGCCGCGAGCATCACAC 7860
|||||
Db 7801 CAGCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCGCAGCCGCGAGCATCACAC 7860
QY 7861 CGCGGAGGCTCAGTGTCTCACTTTCCAGAGTGAAGATGAAGGCATGAAGGAGCTGC 7920
|||||
Db 7861 CGCGGAGGCTCAGTGTCTCACTTTCCAGAGTGAAGATGAAGGCATGAAGGAGCTGC 7920
QY 7921 TGGTGCCACCAAGATCACTCGAGCGCATCAAGCTGCACTCAGGCACAGTGCAGAG 7980
|||||
Db 7921 TGGTGCCACCAAGATCACTCGAGCGCATCAAGCTGCACTCAGGCACAGTGCAGAG 7980
QY 7981 TGCAGATGAAGAGCAGAAAGTGTCAACCCCTAGTACTACACTGTCTTCTCAAGC 8040
|||||
Db 7981 TGCAGATGAAGAGCAGAAAGTGTCAACCCCTAGTACTACACTGTCTTCTCAAGC 8040
QY 8041 GGCAGGCAAGGCTCTGAACCTACTGGGACTTCGGACCGCTTGGGAGCCAGGCTC 8100
|||||
Db 8041 GGCAGGCAAGGCTCTGAACCTACTGGGACTTCGGACCGCTTGGGAGCCAGGCTC 8100
QY 8101 CGCCTTAGTCCCACTCTGAGCCCATGTCTGCCCCCAGCCCAAGGGGACAGGCGCTC 8160
|||||
Db 8101 CGCCTTAGTCCCACTCTGAGCCCATGTCTGCCCCCAGCCCAAGGGGACAGGCGCTC 8160
QY 8161 ACCTTACCAACCCCTAGGTTCCGGTCCCGAGTACAGTGTATCAAAACCCAGATTT 8220
|||||
Db 8161 ACCTTACCAACCCCTAGGTTCCGGTCCCGAGTACAGTGTATCAAAACCCAGATTT 8220
QY 8221 TCTCAGCTCAAGACCCAGGCTCTGCCCGCTGTTAGAAATATAGTCTCTTCCCGAG 8280
|||||
Db 8221 TCTCAGCTCAAGACCCAGGCTCTGCCCGCTGTTAGAAATATAGTCTCTTCCCGAG 8280
QY 8281 AATCCAGCGCGCCAAATGGAACCTCACGCTGGGTCTTAATACCAGTCTTTAAAGGCC 8340
|||||
Db 8281 AATCCAGCGCGCCAAATGGAACCTCACGCTGGGTCTTAATACCAGTCTTTAAAGGCC 8340
QY 8341 AGCCCTTAGAACCCCAAGCTCTCTCGGAACCGCTACCTAGAGCCAGCAACGTTAC 8400
|||||
Db 8341 AGCCCTTAGAACCCCAAGCTCTCTCGGAACCGCTACCTAGAGCCAGCAACGTTAC 8400
QY 8401 TCAGGCTCTCCAGCTGTAGGAGCTGAGGTTTCAACCTTAACCCAGGAGCACAGG 8460
|||||
Db 8401 TCAGGCTCTCCAGCTGTAGGAGCTGAGGTTTCAACCTTAACCCAGGAGCACAGG 8460

QY 8461 TCCCACCTCCAGCCGGGAGGCTTAGGACCACTCAGCCCTAGGAGTATATTTCCGCACT 8520
|||||
Db 8461 TCCCACCTCCAGCCGGGAGGCTTAGGACCACTCAGCCCTAGGAGTATATTTCCGCACT 8520
QY 8521 TCAGAAATTCATATCTTGGAAATCCAAAGTCCCTGCCCCAAATTAATCTCAGTCTGCTTC 8580
|||||
Db 8521 TCAGAAATTCATATCTTGGAAATCCAAAGTCCCTGCCCCAAATTAATCTCAGTCTGCTTC 8580
QY 8581 CAGAAATTTGGAATCCTAGTTCTCTCTCTGTTATCCGAGTCTGGGACACAAACTCC 8640
|||||
Db 8581 CAGAAATTTGGAATCCTAGTTCTCTCTCTGTTATCCGAGTCTGGGACACAAACTCC 8640
QY 8641 GCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAACTGGCCCCGGATC 8700
|||||
Db 8641 GCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTGACGAACTGGCCCCGGATC 8700
QY 8701 AGAGCAGGACCTCCCTTCGGACCCCTTGGGAACCTCCCAAGAGTCCAGCCCATCTCGGAG 8760
|||||
Db 8701 AGAGCAGGACCTCCCTTCGGACCCCTTGGGAACCTCCCAAGAGTCCAGCCCATCTCGGAG 8760
QY 8761 CATCCCGAGGAAATCTGCAGAGGGTTAGGAGTGGGTGACAAGACCTGATCTCTTCT 8820
|||||
Db 8761 CATCCCGAGGAAATCTGCAGAGGGTTAGGAGTGGGTGACAAGACCTGATCTCTTCT 8820
QY 8821 GTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAAGATGATATATTTGTTAAAAAA 8880
|||||
Db 8821 GTTTGTACATAGATTTATTTTTCAGTTTCCAAAGAAAGATGATATATTTGTTAAAAAA 8880
QY 8881 ATATAAGCGCAAGTCCATGTTTATCTGGGAAATTTGGGATGGGCGGAGTGGAGCGC 8940
|||||
Db 8881 ATATAAGCGCAAGTCCATGTTTATCTGGGAAATTTGGGATGGGCGGAGTGGAGCGC 8940
QY 8941 CCTCTCTCCCTTTGCTCTGCTCCCGGACTTTGCGCTCCCTACCTGTGGAGCGCA 9000
|||||
Db 8941 CCTCTCTCCCTTTGCTCTGCTCCCGGACTTTGCGCTCCCTACCTGTGGAGCGCA 9000
QY 9001 GGCAGAGTGGCGGCGGAGGAGTCTAGGCTCCCGCGGCTTGGGCTTCCCGCGCGCG 9060
|||||
Db 9001 GGCAGAGTGGCGGCGGAGGAGTCTAGGCTCCCGCGGCTTGGGCTTCCCGCGCGCG 9060
QY 9061 CCGAGGCGCGTCCCGGCGGCGCTCTCCGAGCTGGGCTGGGCATCCCGGCGCG 9120
|||||
Db 9061 CCGAGGCGCGTCCCGGCGGCGCTCTCCGAGCTGGGCTGGGCATCCCGGCGCG 9120
QY 9121 GCGCGCGCGCGGCTTACGCGCGCGCTTCCGCGGCTTACAGACCGCGCGCGCGCG 9180
|||||
Db 9121 GCGCGCGCGCGGCTTACGCGCGCGCTTCCGCGGCTTACAGACCGCGCGCGCGCG 9180

RESULT 2

AC000134/c
LOCUS AC000134 203300 bp DNA PRI 06-MAY-1999
DEFINITION Homo sapiens Chromosome 11q13 BAC Clone 137c7, complete sequence.
AC000134
ACCESSION AC000134 GI:4755212
VERSION AC000134.14
KEYWORDS HTG.
SOURCE human.

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

1 (bases 1 to 203300)

Crabtree, J. and Roe, B.A.

sequence of a 11q13 bac mapping to PYGM

Unpublished (1997)

REFERENCE

2 (bases 1 to 203300)

Chandrasekharappa, S.C., Guru, S.C., Manickam, P., Olufemi, S.-E. and

Collins, F.S.

Sequence of a 11q13 bac mapping to PYGM

Unpublished (1997)

REFERENCE

3 (bases 1 to 203300)

Emmert-Buck, M.R., Debelenko, L.V., Zhuang, Z., Lubensky, I.A. and

Liotta, L.A.

Sequence of a 11q13 bac mapping to PYGM

JOURNAL REFERENCE	Unpublished (1997)	
AUTHORS	4 (bases 1 to 203300)	
TITLE	Weisemann,J. and Boguski,M.S.	
JOURNAL	Sequence of a 11q13 bac mapping to PYGM	
REFERENCE	Unpublished (1997)	
AUTHORS	5 (bases 1 to 203300)	
TITLE	Agarwal,S.K., Kester,M., Kim,Y.S., Heppner,C., Dong,Q., Spiegel,A.M., Burns,L.A. and Marx,S.J.	
JOURNAL	Sequence of a 11q13 bac mapping to PYGM	
REFERENCE	Unpublished (1997)	
AUTHORS	6 (bases 1 to 203300)	
TITLE	Roe,B.A.	
JOURNAL	Direct Submission Submitted (12-MAR-1997) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
REFERENCE	7 (bases 1 to 203300)	
AUTHORS	Roe,B.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (23-MAR-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
REFERENCE	8 (bases 1 to 203300)	
AUTHORS	Roe,B.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (24-APR-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
REFERENCE	9 (bases 1 to 203300)	
AUTHORS	Roe,B.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (30-APR-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
REFERENCE	10 (bases 1 to 203300)	
AUTHORS	Roe,B.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (01-MAY-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
REFERENCE	11 (bases 1 to 203300)	
AUTHORS	Roe,B.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (04-MAY-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
REFERENCE	12 (bases 1 to 203300)	
AUTHORS	Roe,B.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (05-MAY-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
REFERENCE	13 (bases 1 to 203300)	
AUTHORS	Roe,B.A.	
TITLE	Direct Submission	
JOURNAL	Submitted (06-MAY-1999) Department Of Chemistry And Biochemistry, The University Of Oklahoma, 620 Parrington Oval, Room 208, Norman, OK 73019, USA	
COMMENT	On May 6, 1999 this sequence version replaced gi:4753208.	
FEATURES	Location/Qualifiers	
source	1. .203300 /organism="Homo sapiens" /db_xref="taxon:9606" /clone="137c7"	
BASE COUNT	49152 a 57085 c 53982 g 43081 t	
ORIGIN		
Query Match	98.5%;	Score 9042; DB 85; Length 203300;
Best Local Similarity	99.8%;	Pred.No. 0;
Matches 9166; Conservative	0; Mismatches	6; Indels 11; Gaps 11
QY	1 CTGCTCTGAATCTCTGGCCTCAAGCAATCTCTCGTTCAGCTTCCCAAGTGTGTAA	60

[illegible]

10

Db	163006	TGTGTGTGTATATATATATATATATTTATGTATATGCACATACACAAAAATTTAGCGG	16294
Qy	1141	GAGTGGTGGCGACCGCTGTGATCACAGCTACTCGGAGGCTGAGGCACGAGAATCGCTT	1200
Db	162946	GAGTGGTGGCGACCGCTGTGATCACAGCTACTCGGAGGCTGAGGCACGAGAATCGCTT	162887
Qy	1201	GAGCCGTGAAGTCGAGGCTGCAGTGAGCCACAGATCGAGCCACTGCATTCACGCTGGG	1260
Db	162886	GAGCCGTGAAGTCGAGGCTGCAGTGAGCCACAGATCGAGCCACTGCATTCACGCTGGG	162827
Qy	1261	GAAGAAGAGACCGTGTCTCAAAACAAACAAACAAAGCTACTCTTAGCACGTGTAGA	1320
Db	162826	GAAGAAGAGACCGTGTCTCAAAACAAACAAACAAAGCTACTCTTAGCACGTGTAGA	162767
Qy	1321	GTATCTCGCGGGGAAGTGGGAACGAGTCTGCACACAGATAGGCATCTTTATATGT	1380
Db	162766	GTATCTCGCGGGGAAGTGGGAACGAGTCTGCACACAGATAGGCATCTTTATATGT	162707
Qy	1381	TACACAGACTGATACCAGCTAAAGCGCTGAACACATTTACTCTCTGGCAGTGTTTAA	1440
Db	162706	TACACAGACTGATACCAGCTAAAGCGCTGAACACATTTACTCTCTGGCAGTGTTTAA	162647
Qy	1441	AAGTATCTGTTTTCTCATATATGTTTTATTTTAAATTTTTCTGATCAAGCAACCTGATC	1500
Db	162646	AAGTATCTGTTTTCTCATATATGTTTTATTTTAAATTTTTCTGATCAAGCAACCTGATC	162587
Qy	1501	TTTTTCTCATAACTTGCACACGACCCGCTGACAGCAAAACCGCAGAGCTCGCGGACC	1560
Db	162586	TTTTTCTCATAACTTGCACACGACCCGCTGACAGCAAAACCGCAGAGCTCGCGGACC	162527
Qy	1561	TCCACCCCGAGTCTGCAGTAGTGCCCCCGGACTACATTTTCCAGAGCACTTCGGGG	1620
Db	162526	TCCACCCCGAGTCTGCAGTAGTGCCCCCGGACTACATTTTCCAGAGCACTTCGGGG	162467
Qy	1621	ACGCTTCTCGCTCGCTCGGCTGGAAGGAAGGCCAATCCTCAGTATCTCTCGGAAGGAG	1680
Db	162466	ACGCTTCTCGCTCGCTCGGCTGGAAGGAAGGCCAATCCTCAGTATCTCTCGGAAGGAG	162407
Qy	1681	GTGTCGGAGCGCGGACCTAGAGATCCCAAGACCACAGCGAGGGCCCGCCGCCCA	1740
Db	162406	GTGTCGGAGCGCGGACCTAGAGATCCCAAGACCACAGCGAGGGCCCGCCGCCCA	162347
Qy	1741	CTATTTCCAGGCTCTGGGGGCGAGGGTGGGCCACAGACTCCAATTTCCCGCGGGTAgTGC	1800
Db	162346	CTATTTCCAGGCTCTGGGGGCGAGGGTGGGCCACAGACTCCAATTTCCCGCGGGTAgTGC	162287
Qy	1801	GACCTAGGGCGGGACTTCATGTCCACAGAGGCTCCGGGCGGCTGCCCGGGTGCCT	1860
Db	162286	GACCTAGGGCGGGACTTCATGTCCACAGAGGCTCCGGGCGGCTGCCCGGGTGCCT	162227
Qy	1861	AGTGTGGGATTAAGCGCGAGTGGCGAGGGGACCGAGGCCAGGACTCTCTTTGGGG	1920
Db	162226	AGTGTGGGATTAAGCGCGAGTGGCGAGGGGACCGAGGGCCAGGACTCTCTTTGGGG	162167
Qy	1921	TTTGGGGCTTGACCTGGGTGCGCTTTCTGCACAGACTTTTACAGCCCCCGGGGCAAGT	1980
Db	162166	TTTGGGGCTTGACCTGGGTGCGCTTTCTGCACAGACTTTTACAG - CCCC GGGGCAAGT	162108
Qy	1981	CGTAGAGAGGGGCGGGCGGCCATTGGGGCTCTCATTTGGGGTGTCTGGGGCGACCCC	2040
Db	162107	CGTAGAGAGGGGCGGGCGGCCATTGGGGCTCTCATTTGGGGTGTCTGGGGCGACCCC	162048
Qy	2041	ATCGGGTACCGGGCGTCCCGGAATTTGGGGGACAAAAGGCTCTGCAGTCTCGCTGAG	2100
Db	162047	ATCGGGTACCGGGCGTCCCGGAATTTGGGGGACAAAAGGCTCTGCAGTCTCGCTGAG	161988
Qy	2101	GGGTCTCACCGACAAAAG - AGGGGAACCGGTGAGCAGAGGCTGAAGAGGGTGGGAAGC	2159
Db	161987	GGGTCTCACCGACAAAAGAGGGGAACCGGTGAGCAGAGGCTGAAGAGGGTGGGAAGC	161928
Qy	2160	AGGGAGGCTGTGCTGTGTGCGGGCGGGTGGAACTTTAGCGGACCTTAGGAGGAGGCTCC	2219
Db	161927	AGGGAGGCTGTGCTGTGTGCGGGCGGGTGGAACTTTAGCGGACCTTAGGAGGAGGCTCC	161868

QY	2220	CGGCGCGAACCTGCCGACACCTCCCTCCCGGGTTGCCTTGAGGCGCGCGCCACCGC	2279
Db	161867	CGGCGCGAACCTGCCGACACCTCCCTCCCGGGTTGCCTTGAGGCGCGCGCCACCGC	161808
QY	2280	CGCGCGCATGGGCTGAAGCGCGCCAGAAAGAGCGTGTTCGCGCTGCCTCCATCGAGC	2339
Db	161807	CGCGCGCATGGGCTGAAGCGCGCCAGAAAGAGCGTGTTCGCGCTGCCTCCATCGAGC	161748
QY	2340	ACGTGGTGGCGCTGTGTGCTGCGAGCTGGCGCGAGGAGCGCGACCTGGTGCTGCTTTT	2399
Db	161747	ACGTGGTGGCGCTGTGTGCTGCGAGCTGGCGCGAGGAGCGCGACCTGGTGCTGCTTTT	161688
QY	2400	CCTTGGTCTGGGCTTCGTGGAGCATTTCTTGCTGTCAACCGCGTCATCCCTACCAACG	2459
Db	161687	CCTTGGTCTGGGCTTCGTGGAGCATTTCTTGCTGTCAACCGCGTCATCCCTACCAACG	161628
QY	2460	TTCCCGAGCTCACCTTCCAGCCACGCCCGCCCGACCGCTGCGGCTCACCTACT	2519
Db	161627	TTCCCGAGCTCACCTTCCAGCCACGCCCGCCCGACCGCTGCGGCTCACCTACT	161568
QY	2520	TTCCCGTGGCGACCTGTCTATCATTCGCGCGCTCTATGCGCCGCTTCACGCGCCAGATCC	2579
Db	161567	TTCCCGTGGCGACCTGTCTATCATTCGCGCGCTCTATGCGCCGCTTCACGCGCCAGATCC	161508
QY	2580	GAGGCGCGCTCGACCTGTCCCTCTATCTCGAGAAGGGGTGTCTCCAGCGGTGAGCTGG	2639
Db	161507	GAGGCGCGCTCGACCTGTCCCTCTATCTCGAGAAAGGGGTGTCTCCAGCGGTGAGCTGG	161448
QY	2640	TGAAGAAGGTCTCGATGTCTATGGAACAGCCTCAGCGCTCCTACTTCAAGATCGGG	2699
Db	161447	TGAAGAAGGTCTCGATGTCTATGGAACAGCCTCAGCGCTCCTACTTCAAGATCGGG	161388
QY	2700	CCACATCCAGTCCCTCTTTCAGCTTTCATCAGAGTTTGAGGCCAGTAGTGGGAATCTTA	2759
Db	161387	CCACATCCAGTCCCTCTTTCAGCTTTCATCAGAGTTTGAGGCCAGTAGTGGGAATCTTA	161328
QY	2760	TCCATGACCCACTTCTTCAAAACCTCCATGGTTTACAGAACCTTTTAAAGACTCTAAG	2819
Db	161327	TCCATGACCCACTTCTTCAAAACCTCCATGGTTTACAGAACCTTTTAAAGACTCTAAG	161268
QY	2820	CCTTCTGAGGTTTCGCGAGGTGTTATTTTCCTCTTTCAGTTTGGAACCTGAAGCCAGAG	2879
Db	161267	CCTTCTGAGGTTTCGCGAGGTGTTATTTTCCTCTTTCAGTTTGGAACCTGAAGCCAGAG	161208
QY	2880	AGGGAAATGATATGCCAAAGTCACACAGCGATGCGAGGCTTGAAGTGAAGCCTGATC	2939
Db	161207	AGGGAAATGATATGCCAAAGTCACACAGCGATGCGAGGCTTGAAGTGAAGCCTGATC	161148
QY	2940	ACTTGGCTCCAAATCATCAACCTCACCTCTGCCCTTCAGCACCCGCCCTTGGCAGCTG	2999
Db	161147	ACTTGGCTCCAAATCATCAACCTCACCTCTGCCCTTCAGCACCCGCCCTTGGCAGCTG	161088
QY	3000	AACAGCTACAGGAGTTCTTAAGCATGAGACACAGAGGGCGGACAGATTTAGGGGGCAAG	3059
Db	161087	AACAGCTACAGGAGTTCTTAAGCATGAGACACAGAGGGCGGACAGATTTAGGGGGCAAG	161028
QY	3060	AAAGTAAATTTGGGTGCATTTGAGGCGATTAAACAAAAATAATGGCTATGAAGATTTTTT	3119
Db	161027	AAAGTAAATTTGGGTGCATTTGAGGCGATTAAACAAAAATAATGGCTATGAAGATTTTTT	160968
QY	3120	TTTTTTTTTTTTTTTTTGACAGAGGCTTCACCTCTGCTCCCGAGGCTGGAGTGCAGTGTCT	3179
Db	160967	TTTTTTTTTTTTTTTTTGACAGAGGCTTCACCTCTGCTCCCGAGGCTGGAGTGCAGTGTCT	160908
QY	3180	GATCATGGCTCAGCTGCAGCCTTCAGTCTCCCTGGGCTCAGAGATCTCCCAACCTCAGCCTC	3239
Db	160907	GATCATGGCTCAGCTGCAGCCTTCAGTCTCCCTGGGCTCAGAGATCTCCCAACCTCAGCCTC	160848
QY	3240	CTGAGTACTGAGAGTACAGGCATGCACCGTGGTGTGTTAATTTTTGATTTTTTTTTT	3299
Db	160847	CTGAGTACTGAGAGTACAGGCATGCACCGTGGTGTGTTAATTTTTGATTTTTTTTTT	160788

QY 3300 GTAGATGGTGTCTCACTATGTGGCCAGACTGGTCTTGAACCTCTTGGCTCAAGTGAT 3359
Db 160787 GTAGATGGTGTCTCACTATGTGGCCAGACTGGTCTTGAACCTCTTGGCTCAAGTGAT 160728
QY 3360 CTGCCGCCCTCAGTCTCCCAAAATGCTGGGATACAGGTGTAGCCACCGCAACTGGTGGC 3419
Db 160727 CTGCCGCCCTCAGTCTCCCAAAATGCTGGGATACAGGTGTAGCCACCGCAACTGGTGGC 160668
QY 3420 CTATGAAAATTTTTTTTTTTTTCAGACGGCTCTCACTCTGTGCGCCAGGCTGGAGTGC 3479
Db 160667 CTATGAAAATTTTTTTTTTTTTCAGACGGCTCTCACTCTGTGCGCCAGGCTGGAGTGC 160608
QY 3480 AGRGGTCAATCTCGGCTCACTCAAGCTCTGGCTCTCTGCTTTTCATGCCATTTCTCCTGCC 3539
Db 160607 AGRGGTCAATCTCGGCTCACTCAAGCTCTGGCTCTCTGCTTTTCATGCCATTTCTCCTGCC 160548
QY 3540 TCCTGCCCTCAGCCTCTGAGTGTGGGACTACAGAGCCTGCCACCATGCCCTGGCTAAT 3599
Db 160547 TCCTGCCCTCAGCCTCTGAGTGTGGGACTACAGAGCCTGCCACCATGCCCTGGCTAAT 160488
QY 3600 TTTTTTTTTGGATTTTGTAGTAGACGAGGTTTCACCAATGTTAGCCAGATGGTCTCGATC 3659
Db 160487 TTTTTTTTTGGATTTTGTAGTAGACGAGGTTTTCACCAATGTTAGCCAGATGGTCTCGATC 160428
QY 3660 TCCTGACCTCGTGATCGCGCCGCTTGGCCTGCCAAAGTGTGGGATTCAGACGCTGAGC 3719
Db 160427 TCCTGACCTCGTGATCGCGCCGCTTGGCCTGCCAAAGTGTGGGATTCAGACGCTGAGC 160368
QY 3720 CACGGACCTGGTCAAAAATGTTTGACAGACAGAGGGGCTTGACCTCAAAAGGCTTAAG 3779
Db 160367 CACGGACCTGGTCAAAAATGTTTGACAGACAGAGGGGCTTGACCTCAAAAGGCTTAAG 160308
QY 3780 AGTCAGGGCTTGCAAGAGCTTTTGACAAAGCCGGTTGACTGGCAATCCATCTCGTG 3839
Db 160307 AGTCAGGGCTTGCAAGAGCTTTTGACAAAGCCGGTTGACTGGCAATCCATCTCGTG 160248
QY 3840 TGGCATATCAGAAAGGAATCAGAGGCTGCTTCTCAGCTTAGCAGGAAAGAGTGCAGAGA 3899
Db 160247 TGGCATATCAGAAAGGAATCAGAGGCTGCTTCTCAGCTTAGCAGGAAAGAGTGCAGAGA 160188
QY 3900 TAAATAGGGGTATTTCTGCTGGTGTATAGCCAGAGGTGTGGCCAGCGTCTCTGTTT 3959
Db 160187 TAAATAGGGGTATTTCTGCTGGTGTATAGCCAGAGGTGTGGCCAGCGTCTCTGTTT 160128
QY 3960 TTGCCATTCCTGTTTTAACCTAGTAAGTGCAGTAAATGGAATCCCTTAAATCCATAGAAT 4019
Db 160127 TTGCCATTCCTGTTTTAACCTAGTAAGTGCAGTAAATGGAATCCCTTAAATCCATAGAAT 160068
QY 4020 ATATAATAGAGTTGCAGAGAAAGCAGAGGTAGGGCCAAAGGCTGGGTACAGGATA 4079
Db 160067 ATATAATAGAGTTGCAGAGAAAGCAGAGGTAGGGCCAAAGGCTGGGTACAGGATA 160008
QY 4080 TCCAGAAAGGTATCTCTGTGGACATACAGGGTGTAAACAGGGAGAGTCTTTGAACACG 4139
Db 160007 TCCAGAAAGGTATCTCTGTGGACATACAGGGTGTAAACAGGGAGAGTCTTTGAACACG 159948
QY 4140 TGGGAGGAAGGATAGTGGGAGGATAGTGGGAGGAAATCTCAGGTTGGGTACAGGCTTG 4199
Db 159947 TGGGAGGAAGGATAGTGGGAGGATAGTGGGAGGAAATCTCAGGTTGGGTACAGGCTTG 159888
QY 4200 GAAAGGAGTGGGAGGAGTGTGCCCATCACTACCTGG-CCCCTTTCCCATGTTAAAG 4258
Db 159887 GAAAGGAGTGGGAGGAGTGTGCCCATCACTACCTGGCCCCCTTTTCCCATGTTAAAG 159828
QY 4259 CACAGAGACCTCTTTTCACTTACCTCCCTTCCACAGGACCAAAATGGACAGCTCCGG 4318
Db 159827 CACAGAGACCTCTTTTCACTTACCTCCCTTCCACAGGACCAAAATGGACAGCTCCGG 159768
QY 4319 TGTGGCCTTTTCTGTGGTGTGGGCCCTCCAGGCCCTGGGTCTCGGGATGTCCACCTCGC 4378
Db 159767 TGTGGCCTTTTCTGTGGTGTGGGCCCTCCAGGCCCTGGGTCTCGGGATGTCCACCTCGC 159708
QY 4379 CCTGTCTGAGGATCATGCCCTGGGTAGTGTGTTGGGCCCCAATGGGGAGCAGACAGCTGAGGT 4438

Db 159707 CCTGTCTGAGGATCATGCCCTGGGTAGTGTGTTGGGCCCAATGGGAGCAGACAGCTGAGGT 159648
QY 4439 CACCTGGCAGGCAAGGGCAACGAGGACCCAGGGGCCACACAGTCAATGCCGTGTGGC 4498
Db 159647 CACCTGGCAGGCAAGGGCAACGAGGACCCAGGGGCCACACAGTCAATGCCGTGTGGC 159588
QY 4499 TGAGCGGTATTGTTCCCTCCCGCCAGCCTTGTCCCTTTCATCTGTGTAGTCCCAAGCC 4558
Db 159587 TGAGCGGTATTGTTCCCTCCCGCCAGCCTTGTCCCTTTCATCTGTGTAGTCCCAAGCC 159528
QY 4559 ACCCAAGGACTCCATTTTCTTGGGCCACACCCCTTTCCTCCATCACCCACATPAGG 4618
Db 159527 ACCCAAGGACTCCATTTTCTTGGGCCACACCCCTTTCCTCCATCACCCACATPAGG 159468
QY 4619 AAGGGAAGACAGAGAGCCCTTTTCTGGCTGTCTATTTCCCTGAAGCAGCAGAGGTGG 4678
Db 159467 AAGGGAAGACAGAGAGCCCTTTTCTGGCTGTCTATTTCCCTGAAGCAGCAGAGGTGG 159408
QY 4679 GCCATCATGAGACATATGATCTCATCCCCCCTTAAGAGCTGGCTGTACCTGAAAGGATC 4738
Db 159407 GCCATCATGAGACATATGATCTCATCCCCCCTTAAGAGCTGGCTGTACCTGAAAGGATC 159348
QY 4739 ATACATGCGCTGTGACGCAAGATGGAGTGGGCTTTCATGGTGTGCCATCAACCCCTTC 4798
Db 159347 ATACATGCGCTGTGACGCAAGATGGAGTGGGCTTTCATGGTGTGCCATCAACCCCTTC 159288
QY 4799 CATTTGACCTGCACACGACTCGCTGGAGCTTCTGACGCTGCACGAGGTGAGGCTGAGCC 4858
Db 159287 CATTTGACCTGCACACGACTCGCTGGAGCTTCTGACGCTGCACGAGGTGAGGCTGAGCC 159228
QY 4859 AATGGGCGAGGACTGGGCTAGGCGCAGACTTGTACTTGTGTGGGACCCCTGGCGAGGSCAC 4918
Db 159227 AATGGGCGAGGACTGGGCTAGGCGCAGACTTGTACTTGTGTGGGACCCCTGGCGAGGSCAC 159168
QY 4919 TTTTCCCTTCTGAGCTTCAGCTTCCCTCTGGAAAATGGGTTAGTAAATTTCTTGGCCTG 4978
Db 159167 TTTTCCCTTCTGAGCTTCAGCTTCCCTCTGGAAAATGGGTTAGTAAATTTCTTGGCCTG 159108
QY 4979 GCCTTTCCAGGGCTCTTGGGAGAGTAGAATTGAGATGTGAAATTTGCTTGACTTCCATTA 5038
Db 159107 GCCTTTCCAGGGCTCTTGGGAGAGTAGAATTGAGATGTGAAATTTGCTTGACTTCCATTA 159048
QY 5039 AAGGCTGTGTCAGAAATTTTGGCCCTTCCACATGGTGGGTGGTCCCTGTGTTGTTCTGAC 5098
Db 159047 AAGGCTGTGTCAGAAATTTTGGCCCTTCCACATGGTGGGTGGTCCCTGTGTTGTTCTGAC 158988
QY 5099 CCCCACCTCTGCCGATAGGCTAAGGACCCGTTCTCCTCCTGTTCCGTGGCTCATAACT 5158
Db 158987 CCCCACCTCTGCCGATAGGCTAAGGACCCGTTCTCCTCCTGTTCCGTGGCTCATAACT 158928
QY 5159 CTCTCTCTTCCGCTCCTAAGAGCTGCTCTGGCTGCTCTATGACCTTGGGACATCTGGAAGG 5218
Db 158927 CTCTCTCTTCCGCTCCTAAGAGCTGCTCTGGCTGCTCTATGACCTTGGGACATCTGGAAGG 158868
QY 5219 TCAGTAGAGGAAGTGGCCAGGCTGCGCCTTGGTGTAGGCGGGGGCTGGGTGGCAGCCTG 5278
Db 158867 TCAGTAGAGGAAGTGGCCAGGCTGCGCCTTGGTGTAGGCGGGGGCTGGGTGGCAGCCTG 158808
QY 5279 AATTATCATCTTTCTCCTAGGTACCCCATGGCTTAGGGAACCTTGGCAGATCTAGAGGAGC 5338
Db 158807 AATTATCATCTTTCTCCTAGGTACCCCATGGCTTAGGGAACCTTGGCAGATCTAGAGGAGC 158748
QY 5339 TGGAGCCACCCCTTGGCGGGCCAGACCCACTCACCTCTACCAAGGTGGGGCATCTA 5398
Db 158747 TGGAGCCACCCCTTGGCGGGCCAGACCCACTCACCTCTACCAAGGTGGGGCATCTA 158688
QY 5399 AGGAGGTGAGAAAGGAGAGCCCTTAACAGTGGCTGAGGAGGGGCCCTCATCTGGGGAGA 5458
Db 158687 AGGAGGTGAGAAAGGAGAGCCCTTAACAGTGGCTGAGGAGGGGCCCTCATCTGGGGAGA 158628
QY 5459 TGAGAACAGAACTTTTGTGTGTTGGGGGTATCGGCCCATCCAGTCTCACTTTGTGTCAACT 5518

Db 158627 TGAGAGAGAACTTTGTGTGTGGGGGTATCGCCCATCCAGTCTCACATTTTGTGTCAACT 158568
QY 5519 GTGTGCAGAAATCAGTTTCACTCAGGGCTGTCTCAGGGGTGTCCAGGGTTCCTCCAGCCTGGG 5578
Db 158567 GTGTGCAGAAATCAGTTTCACTCAGGGCTGTCTCAGGGGTGTCCAGGGTTCCTCCAGCCTGGG 158508
QY 5579 AGTGCAGGGGTGCAATTTGTGTCCTCAGCCCTCAGCCCTGCTTTTCTGCCACTGCTTACTGTCTCT 5638
Db 158507 AGTGCAGGGGTGCAATTTGTGTCCTCAGCCCTCAGCCCTGCTTTTCTGCCACTGCTTACTGTCTCT 158448
QY 5639 TCCTGGAGTATTAACAGAGGTCAAAATGTGGTAGAGCACTGAAGAGGGGTGTTCACTTGG 5698
Db 158447 TCCTGGAGTATTAACAGAGGTCAAAATGTGGTAGAGCACTGAAGAGGGGTGTTCACTTGG 158388
QY 5699 TGGGTGTAGTGGGAGAGGGCCATTTGGGCTGGGCTTGAAGTCTTTTGGTGATGTGTAG 5758
Db 158387 TGGGTGTAGTGGGAGAGGGCCATTTGGGCTGGGCTTGAAGTCTTTTGGTGATGTGTAG 158328
QY 5759 AAGAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGGGCCAGGCCCCACCCCTGCAGTC 5818
Db 158327 AAGAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGGGCCAGGCCCCACCCCTGCAGTC 158268
QY 5819 TCGCCAGGCCCTCAGCCAGCAGTCTGTAGACCCAGGGAGGAGACAGGTAGAAGGGCTG 5878
Db 158267 TCGCCAGGCCCTCAGCCAGCAGTCTGTAGACCCAGGGAGGAGACAGGTAGAAGGGCTG 158208
QY 5879 GCAGCGAGTGGAGTGGGAGTGAGATGGAGAGGACTCCCTGGGATCTTCCCTGTGGCCCC 5938
Db 158207 GCAGCGAGTGGAGTGGGAGTGAGATGGAGAGGACTCCCTGGGATCTTCCCTGTGGCCCC 158148
QY 5939 TTCTGGGTGTGCCCTGGTGGGGCATTTGTGCCAGAGGGCAGCTGGGGTGCCTCCCTGA 5998
Db 158147 TTCTGGGTGTGCCCTGGTGGGGCATTTGTGCCAGAGGGCAGCTGGGGTGCCTCCCTGA 158088
QY 5999 GGATCCTGTGCTCACCCTCACCCTCAGGGCATTTGCCCTCAGCCAGGACCTACTATCGGGATG 6058
Db 158087 GGATCCTGTGCTCACCCTCACCCTCAGGGCATTTGCCCTCAGCCAGGACCTACTATCGGGATG 158028
QY 6059 AACACATCTACCCCTACATGTACCTGGCTGGCTACACCTGTCGCAACCCCAATGTGCGGG 6118
Db 158027 AACACATCTACCCCTACATGTACCTGGCTGGCTACACCTGTCGCAACCCCAATGTGCGGG 157968
QY 6119 AAGCCCTCGAGCCCTGGCGGACAGGGCCACTGTCTATCCAGGAGTGAGGATCCCCCTACT 6178
Db 157967 AAGCCCTCGAGCCCTGGCGGACAGGGCCACTGTCTATCCAGGAGTGAGGATCCCCCTACT 157908
QY 6179 AGGGCTCGAGCCTGTCTTTCTTCTCCCTCCATCAGTTTCCAAACCCCTCTGTCCAGGAC 6238
Db 157907 AGGGCTCGAGCCTGTCTTTCTTCTCCCTCCATCAGTTTCCAAACCCCTCTGTCCAGGAC 157848
QY 6239 TGAGGCTGGCTCCAGCCGCCATCCCTTTCCATCCAGTCCCTAGGCAGCAAGGCCACC 6298
Db 157847 TGAGGCTGGCTCCAGCCGCCATCCCTTTCCATCCAGTCCCTAGGCAGCAAGGCCACC 157788
QY 6299 ATTACCCAGGAGTGGGACCCCTGATTAAGGTGTACATCTTTTCCCTCCCTCCCTCTCC 6358
Db 157787 ATTACCCAGGAGTGGGACCCCTGATTAAGGTGTACATCTTTTCCCTCCCTCCCTCTCC 157728
QY 6359 TCCTAAATTTTTTTTCTCAGAACAGTCTCAATCTCCAATGTTTAAACCAATCATCCA 6418
Db 157727 TCCTAAATTTTTTTTCTCAGAACAGTCTCAATCTCCAATGTTTAAACCAATCATCCA 157668
QY 6419 GCAGTGGGACTTCCACCCCTCGGCCCATCCCTTCCCTCCTTCTTCTTCTTCTCTCTG 6478
Db 157667 GCAGTGGGACTTCCACCCCTCGGCCCATCCCTTCCCTCCTTCTTCTTCTTCTTCTCTG 157608
QY 6479 GCGTACCCAGACAGCATATTTTGCAGTGGAGACCCCACTACTCCCCAGCCCTCGGG 6538
Db 157607 GCGTACCCAGACAGCATATTTTGCAGTGGAGACCCCACTACTCCCCAGCCCTCGGG 157548
QY 6539 GCGTCAATCCCGCAGGTCCTCGGGGTACCCCGGATGGTGAGACCCCTTCAGACCCCT 6598
Db 157547 GCGTCAATCCCGCAGGTCCTCGGGGTACCCCGGATGGTGAGACCCCTTCAGACCCCT 157488

QY 6599 ACAGAGACCCCACTGCTCTCACAGCTTACAACTACTGCCGGGAAGACGAGAGATCTACAA 6658
Db 157487 ACAGAGACCCCACTGCTCTCACAGCTTACAACTACTGCCGGGAAGACGAGAGATCTACAA 157428
QY 6659 GGAGTCTTTTGAAGTAGCAATGATGTATCCCCAACCTGCTGAAGAGGAGCCAGCTT 6718
Db 157427 GGAGTCTTTTGAAGTAGCAATGATGTATCCCCAACCTGCTGAAGAGGAGCCAGCTT 157368
QY 6719 GCTGAGGCGGCGAGGAGCGCCGGGAGCAAAAGCCAGGTCAAAAGGCTGAGGCTCCAG 6778
Db 157367 GCTGAGGCGGCGAGGAGCGCCGGGAGCAAAAGCCAGGTCAAAAGGCTGAGGCTCCAG 157308
QY 6779 CTTGTCTCAGCCTCCCACTGGACAGGCTCCCTTCCACAGGGCCATGGGGCTGCATG 6838
Db 157307 CTTGTCTCAGCCTCCCACTGGACAGGCTCCCTTCCACAGGGCTGCATG 157249
QY 6839 TACGGGATTAGGATGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCTTTTG 6898
Db 157248 TACGGGATTAGGATGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCTTTTG 157189
QY 6899 CTATAACTCAGGTCTCGGCCCCACGTTGGACGGGACTGAAGGTATTTTAGAGTTCCTAC 6958
Db 157188 CTATAACTCAGGTCTCGGCCCCACGTTGGACGGGACTGAAGGTATTTTAGAGTTCCTAC 157129
QY 6959 CCTGTGCCCTTCACTAGTTTTCATGGCCAGACTCCCTCCCTCAGCTCAGGGGTGGAGTAGGAT 7018
Db 157128 CCTGTGCCCTTCACTAGTTTTCATGGCCAGACTCCCTCCCTCAGCTCAGGGGTGGAGTAGGAT 157069
QY 7019 GGTAGCTCCTGCTATGAGATGGCTTTTAAAGAAAGAGGTTCCTAAGAAATGTTCCCAA 7078
Db 157068 GGTAGCTCCTGCTATGAGATGGCTTTTAAAGAAAGAGGTTCCTAAGAAATGTTCCCAA 157009
QY 7079 CCTATGCTTACCTTTCTGAGCCAGGGCTCTTGCCTAGGTGGGGGCTTGCCCTGTGC 7138
Db 157008 CCTATGCTTACCTTTCTGAGCCAGGGCTCTTGCCTAGGTGGGGGCTTGCCCTGTGC 156949
QY 7139 CCTCTGCTAAGGGGTGAGTAAGAGACTGATCTGTGCCCTCCCTTCCCTTCCGTCACAGGC 7198
Db 156948 CCTCTGCTAAGGGGTGAGTAAGAGACTGATCTGTGCCCTCCCTTCCCTTCCGTCACAGGC 156889
QY 7199 ACCAGAGCAAGGTTCCGCTCCAGGACCTGAGTGTTCGCCCACTGTCGCGGATTC 7258
Db 156888 ACCAGAGCAAGGTTCCGCTCCAGGACCTGAGTGTTCGCCCACTGTCGCGGATTC 156829
QY 7259 TAGCAGGATCTGCAAAATGGGAGGAGGAGCTCCCAACCCCTGCTGCACGTCGGCTGG 7318
Db 156828 TAGCAGGATCTGCAAAATGGGAGGAGGAGCTCCCAACCCCTGCTGCACGTCGGCTGG 156769
QY 7319 GGCACCTTTTGTGAGTCCCTTAGGCCCTTTTGAAGGACAGGTGAGGACAGCTGCACA 7378
Db 156768 GGCACCTTTTGTGAGTCCCTTAGGCCCTTTTGAAGGACAGGTGAGGACAGCTGCACA 156709
QY 7379 GAGGTCTGGCACTACAGTGTGTGACAGGACGCGGCTGTGTACAGCTTTTCTTGCCCA 7438
Db 156708 GAGGTCTGGCACTACAGTGTGTGACAGGACGCGGCTGTGTACAGCTTTTCTTGCCCA 156649
QY 7439 GGGGAGCATCTGCCATCCCTTCCGTGCCGATGGACTGAGACCCCTGGGTGGGATG 7498
Db 156648 GGGGAGCATCTGCCATCCCTTCCGTGCCGATGGACTGAGACCCCTGGGTGGGATG 156589
QY 7499 GGATGCCAGAGCAGGGTCTCTGGAGTTCCAGCCACTGGCCGGCAACCTTGTCTCACTT 7558
Db 156588 GGATGCCAGAGCAGGGTCTCTGGAGTTCCAGCCACTGGCCGGCAACCTTGTCTCACTT 156529
QY 7559 GGTCTCCCACTGGCCAGGTGCGGCAAGGTGCGCATAGTAGCCGAGAGCCGAGGC 7618
Db 156528 GGTCTCCCACTGGCCAGGTGCGGCAAGGTGCGCATAGTAGCCGAGAGCCGAGGC 156469
QY 7619 GSCCGAGCCGAGGAGCCGTGGGGCGAGGAAGCCCGGGAAGCCCGCGGGCCACG 7678
Db 156468 GSCCGAGCCGAGGAGCCGTGGGGCGAGGAAGCCCGGGAAGCCCGCGGGCCACG 156409

[illegible]

Db	155331	GAGCATCCCGAGGAATCTGCAGA	GGGTTAGAGTGGGTGACAAGACCTGATCTCTTT	155273
Qy	8818	CCTGTTTTGTACATAGATTTATTTT	CAGTTCCAAGAAAGATGAATACATTTTGTTTAAAA	8877
Db	155272	CCTGTTTTGTACATAGATTTATTTT	CAGTTCCAAGAAAGATGAATACATTTTGTTTAAAA	155213
Qy	8878	AAATATAAAGCCCAAGTCCATCTTATCT	GTGCAAAATGGGGATGGGGCGGAGTGGAG	8937
Db	155212	AAATATAAAGCCCAAGTCCATCTTATCT	GTGCAAAATGGGGATGGGGCGGAGTGGAG	155153
Qy	8938	CGCCCTCTTCTCCCTTTTGTCTTCTG	GCTCCCGGAGCTTTTGCCTCCCTACCTGTGGAGCG	8997
Db	155152	CGCCCTCTTCTCCCTTTTGTCTTCTG	GCTCCCGGAGCTTTTGCCTCCCTACCTGTGGAGCG	155093
Qy	8998	CGAGCGACAGTGGCGGGGAAGACGTAG	CGCTCCGCCCGGCTTTGGGGCTTTCCCGCGCG	9057
Db	155092	CGAGCGACAGTGGCGGGGAAGACGTAG	CGCTCCGCCCGGCTTTGGGGCTTTCCCGCGCG	155033
Qy	9058	CCGCCGAGGGCCGCTCCCGCGGGCGCT	CTCTCCCGGACTGGCGGTGGGGCATCCGNGGCG	9117
Db	155032	CCGCCGAGGGCCGCTCCCGCGGGCGCT	CTCTCCCGGACTGGCGGTGGGGCATCCGNGGCG	154974
Qy	9118	CGCGCCCGCGCCCGCGGCTTTCAGCC	CGCCCGCGGCTTTCAGAGCACCGGGCGCCCGCC	9177
Db	154973	CGCGCCCGCGCCCGCGGCTTTCAGCC	CGCCCGCGGCTTTCAGAGCACCGGGCGCCCGCC	154914
Qy	9178	CCG 9180	1111	
Db	154913	CCG 154911		
RESULT	3			
LOCUS	HS838L14	109290 bp	DNA	31-MAR-1998
DEFINITION	Homo sapiens chromosome 11 clone 838L14 map 11q13, *** SEQUENCING IN PROGRESS ***, 15 unordered pieces.			
ACCESSION	Y12335			
VERSION	Y12335.1	GI:2208954		
KEYWORDS	HTG; HTGS_PHASE1.			
SOURCE	human.			
ORGANISM	Homo sapiens			
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
AUTHORS	Kedra,D., Seroussi,E., Fransson,I., Trifunovic,J., Clark,M., Lagercrantz,J., Blennow,E., Mehlin,H. and Dumanski,J.			
TITLE	The germinal center kinase gene and a novel CDC23-like gene are located in the vicinity of the PYGM gene on 11q13			
JOURNAL	Hum. Genet.	100 (5-6),	611-619	(1997)
MEDLINE	98001089			
REFERENCE	2	(bases 1 to 109290)		
AUTHORS	Kedra,D.			
TITLE	Direct Submission			
JOURNAL	Submitted (20-MAR-1997) D. Kedra, Karolinska Hospital, Department Of Molecular Medicine, Building L-6, S-171 76 Stockholm, SWEDEN			
COMMENT	Warning!			
	This sequence is unfinished. It is not contiguous, consists of 161 contigs separated by gaps of unknown length. Gaps in the sequence are represented by a stretch of NNNN. . . . Order of the contigs unknown.			
	1	2421	contig of	2421
	2612	4907	contig of	2296
	5098	5742	contig of	645
	5933	6485	contig of	553
	6676	27557	contig of	20882
	27748	31946	contig of	4199
	32137	40306	contig of	8170
	40497	40866	contig of	370
	41057	49958	contig of	8902
	50149	53510	contig of	3362
	57041	57416	contig of	3716
	57607	58216	contig of	610

Db 78426 |||||GAGCCGTGAAGTGGAGCTGCACTGAGCCAGATCGAGCCACTGCAATCCAGCCTGGGC 78485
Qy 1361 GAAAGAAAGACCGTGTCTCAAAACAAACAAACAAAGCTACTCTTAGCACGTGTTAGA 1320
Db 78486 GAAAGAAAGACCGTGTCTCAAAACAAACAAACAAAGCTACTCTTAGCACGTGTTAGA 78545
Qy 1321 GTATCTCGCGGGCGAAGTGGGAACAGAGTGTCTGCACACAGAGTAGGCATCTTTATATGT 1380
Db 78546 GTATCTCGCGGGCGAAGTGGGAACAGAGTGTCTGCACACAGAGTAGGCATCTTTATATGT 78605
Qy 1381 TAACAGACACTGATACCCAGCTAAGCGGCTGAACACATTTACTCTCTGGCAGTGTATAA 1440
Db 78606 TAACAGACACTGATACCCAGCTAAGCGGCTGAACACATTTACTCTCTGGCAGTGTATAA 78665
Qy 1441 AAGTATCTGTTTTCTCATATGTTGTTTTATTTATTTTCTGGATCAAGCAACCTGATC 1500
Db 78666 AAGTATCTGTTTTCTCATATGTTGTTTTATTTATTTTCTGGATCAAGCAACCTGATC 78725
Qy 1501 TTTTTCCTCATAACTTCCGACCCGACCCGTGACAGCAAAACCGCAGAAAGCTCGCGGACC 1560
Db 78726 TTTTTCCTCATAACTTCCGACCCGACCCGTGACAGCAAAACCGCAGAAAGCTCGCGGACC 78785
Qy 1561 TCCACACCCGAGTCTGACAGTGTGCCCCCGGACTACATTTTCCAGAAAGCACTGCGGGC 1620
Db 78786 TCCACACCCGAGTCTGACAGTGTGCCCCCGGACTACATTTTCCAGAAAGCACTGCGGGC 78845
Qy 1621 ACGCTTCTGCTGCTGGCTGAGGCTGAAGGAAGGGCCAAATCCCTGAGTATCTCGGGAAGGAG 1680
Db 78846 ACGCTTCTGCTGCTGGCTGAGGCTGAAGGAAGGGCCAAATCCCTGAGTATCTCGGGAAGGAG 78905
Qy 1681 GTGTCGGAGCCGCGGACCTAGAGATCCAGAAAGCCAGAGCCGCGGCGCCGCCA 1740
Db 78906 GTGTCGGAGCCGCGGACCTAGAGATCCAGAAAGCCAGAGCCGCGGCGCCGCCA 78965
Qy 1741 CTATTTCCAGCTCTCGGGGACGGGTGGGCCAGACTCCACTTCCGGGGGTAGTGC 1800
Db 78966 CTATTTCCAGCTCTCGGGGACGGGTGGGCCAGACTCCACTTCCGGGGGTAGTGC 79025
Qy 1801 GACCTTAGGGGCGGACTTCAATGTCCAGCAGCTCGGGGCGGCTGCGCCGCGGTGCT 1860
Db 79026 GACCTTAGGGGCGGACTTCAATGTCCAGCAGCTCGGGGCGGCTGCGCCGCGGTGCT 79085
Qy 1861 AGTGTGGATGTAAGCGGGAGTGGGAGGGGACCGAGGCCAGAGACTCTCCTTGGGG 1920
Db 79086 AGTGTGGATGTAAGCGGGAGTGGGAGGGGACCGAGGCCAGAGACTCTCCTTGGGG 79145
Qy 1921 TTTGGGGCTTGACCTGGGTGCGCTTTCTGACAGACTTTTACAGCCCGCGGGGCGACAGT 1980
Db 79146 TTTGGGGCTTGACCTGGGTGCGCTTTCTGACAGACTTTTACAGCCCGCGGGGCGACAGT 79204
Qy 1981 CGTAGAGAGGGGGGGGCGGCCATTTGGGGCTCCTCAATTTGGGGTGTCTTGGGGCGCACCCC 2040
Db 79205 CGTAGAGAGGGGGGGGCGGCCATTTGGGGCTCCTCAATTTGGGGTGTCTTGGGGCGCACCCC 79264
Qy 2041 ATCGGGTACCGGGGCTCCCGAAATGTGGGGGACAAAGGCTCTGAGTCTCGGGTGA 2100
Db 79265 ATCGGGTACCGGGGCTCCCGAAATGTGGGGGACAAAGGCTCTGAGTCTCGGGTGA 79324
Qy 2101 GGGTCTACCCACAAAGAGGGGAAGCCGGTGAACAGGCTGAAGAGGCTGGGGAAGCA 2160
Db 79325 GGGTCTACCCACAAAGAGGGGAAGCCGGTGAACAGGCTGAAGAGGCTGGGGAAGCA 79384
Qy 2161 GGGAGCTGTGCTGTGTCGGGGGGGTGGAACCTTACGGACCCCTGGGAGGAGGCTCCC 2220
Db 79385 GGGAGCTGTGCTGTGTCGGGGGGGTGGAACCTTACGGACCCCTGGGAGGAGGCTCCC 79444
Qy 2221 CGGCGAACTCTCCCGACCCCTCCCTCCCGGGCTTGCCTTGCAGCGCGCCCGCCACCGCC 2280
Db 79445 CGGCGAACTCTCCCGACCCCTCCCTCCCGGGCTTGCCTTGCAGCGCGCCCGCCACCGCC 79504
Qy 2281 CGCGCCATGGGCTGAAGCCCGCCAGAACCGCTGTTCCCGCTGGCTCCATCGAGCA 2340
Db |||||

Db 79505 CGCGTCAATGGGGCTGAAGGCCGCCAGAGACGCTGTTCCCGTGGCTCCATCGACGA 79564
Qy 2341 CGTGTGCGCTGTTGCTGCCAGCTGGGCCGAGAGCGGACCTGTGTCCTCTTTC 2400
Db 79565 CGTGTGCGCTGTTGCTGCCAGCTGGGCCGAGAGCGGACCTGTGTCCTCTTTC 79624
Qy 2401 CTTGGTCTGGGCTCTGTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTACCAAGT 2460
Db 79625 CTTGGTCTGGGCTCTGTGGAGCATTTTCTGGCTGTCAACCGCTCATCCCTACCAAGT 79684
Qy 2461 TCCCGAGCTCACCTTCAGCCAGCCCGCCCGAGCCGCTGGCGGCTCACCTACTT 2520
Db 79685 TCCCGAGCTCACCTTCAGCCAGCCCGCCCGAGCCGCTGGCGGCTCACCTACTT 79744
Qy 2521 TCCCGTGGCGGACCTGTCTATCATCGCGGCTCTATGCGCGCTTCACGCGCCAGATCCG 2580
Db 79745 TCCCGTGGCGGACCTGTCTATCATCGCGGCTCTATGCGCGCTTCACGCGCCAGATCCG 79804
Qy 2581 AGCGCCGCTCGACCTGTCCCTCTATCTCGAAGAGGGGTGTCTCCAGCCGTGAGCTGT 2640
Db 79805 AGCGCCGCTCGACCTGTCCCTCTATCTCGAAGAGGGGTGTCTCCAGCCGTGAGCTGT 79864
Qy 2641 GAAGAGGCTCTCCGATGTCATATGAACAGCTCAGCGGCTCTACTTCAAGGATCGGGC 2700
Db 79865 GAAGAGGCTCTCCGATGTCATATGAACAGCTCAGCGGCTCTACTTCAAGGATCGGGC 79924
Qy 2701 CCACATCCAGTCCCTCTTCACTCATCAGGTTGGAGCCAGTAGTGGGAATCTTAT 2760
Db 79925 CCACATCCAGTCCCTCTTCACTCATCAGGTTGGAGCCAGTAGTGGGAATCTTAT 79984
Qy 2761 CCATGACCCACTTCTTCAAAACCTCATATGTTTACAGAACCTTTTAAAGACTGTAAAG 2820
Db 79985 CCATGACCCACTTCTTCAAAACCTCATATGTTTACAGAACCTTTTAAAGACTGTAAAG 80044
Qy 2821 CTTCTGAGGTTGCGCAGGTGTTATTTCTCTTTCAGTTCGGAACTGAAGCCACAGCA 2880
Db 80045 CTTCTGAGGTTGCGCAGGTGTTATTTCTCTCTTTCAGTTCGGAACTGAAGCCACAGCA 80104
Qy 2881 GGGAAATGATATCCAAAGTCACACAGGCTGGCAGGCTGGAAGCTGAATPCA 2940
Db 80105 GGGAAATGATATCCAAAGTCACACAGGCTGGCAGGCTGGAAGCTGAATPCA 80164
Qy 2941 CTTGGCTCCAAATCATCAACCTCAGCTTGCCTCAGCAGCCACCTTGCACACTGA 3000
Db 80165 CTTGGCTCCAAATCATCAACCTCAGCTTGCCTCAGCA-CCCCACCTTGCACACTGA 80223
Qy 3001 ACAGCTACAGGAGTCTTAAGCATGAGACACAGGGCGGACAGAGTTAGGGGCAAGA 3060
Db 80224 ACAGCTACAGGAGTCTTAAGCATGAGACACAGGGCGGACAGAGTTAGGGGCAAGA 80283
Qy 3061 AGATGAAATTTGGGCTGCATTTGAGGCAAGTTAAACAAATAATGGCTATGAAGATTTT 3120
Db 80284 AGATGAAATTTGGGCTGCATTTGAGGCAAGTTAAACAAATAATGGCTATGAAGATTTT 80343
Qy 3121 TTTTTTTTTTTTTTTGAACAGGGTCTCACTCTGTCCCGCAGGCTGGAGTGCAGTGGTG 3180
Db 80344 TTTTTTTTTTTTTTTGAACAGGGTCTCACTCTGTCCCGCAGGCTGGAGTGCAGTGGTG 80403
Qy 3181 ATCATGGCTCACTCAGCCTCAGTCTCCCTGGGCTCAGAGATCTCCAACTCAGCCTCC 3240
Db 80404 ATCATGGCTCACTCAGCCTCAGTCTCCCTGGGCTCAGAGATCTCCAACTCAGCCTCC 80463
Qy 3241 TGAGTAGCTGAGAGTACAGGATGACCCGGTGTCTGTTTAAATTTTTTGTATTTTTTTG 3300
Db 80464 TGAGTAGCTGAGAGTACAGGATGACCCGGTGTCTGTTTAAATTTTTTGTATTTTTTTG 80523
Qy 3301 TAGAGATGGTGTCTCACTATGTGGCCAGAGTGGTCTTGAACCTTGGGCTCAAGTGTATC 3360
Db 80524 TAGAGATGGTGTCTCACTATGTGGCCAGAGTGGTCTTGAACCTTGGGCTCAAGTGTATC 80583
Qy 3361 TGCCCGCTCAGTCTCCCAAAATGCTGGGATTTACAGGTTGAGCCACCCCAACTGGTGGCC 3420
Db 80584 TGCCCGCTCAGTCTCCCAAAATGCTGGGATTTACAGGTTGAGCCACCCCAACTGGTGGCC 80643

Db 82804 GTGGCAGGGGCTGATTTGTCCTCCCTCAGCCCTGCTTTCTGCCACTGCTTACTGTCCTT 82863
Qy 5640 CCTGGAGTATAACAGAGGTCAAAATGTTGGTAGGACACTGAAGAGGGGTGTTCACTTGGT 5699
Db 82864 CCTGGAGTATAACAGAGGTCAAAATGTTGGTAGGACACTGAAGAGGGGTGTTCACTTGGT 82923
Qy 5700 GGGTGTAGGTGGGAGGAGGCCATTTGGGCTGGGCTTTGAAAGTCTTTGGTGATGTGAGA 5759
Db 82924 GGGTGTAGGTGGGAGGAGGCCATTTGGGCTGGGCTTTGAAAGTCTTTGGTGATGTGAGA 82983
Qy 5760 AGAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGAGGCCAGGCCCCACCCCTCGAGTCT 5819
Db 82984 AGAGTGTCTGAGAAAGAGAGGGCCCTGAGCTCGAGAGGCCAGGCCCCACCCCTCGAGTCT 83043
Qy 5820 GCCCAGGCCCTCAGCCAGCAGCTCTGTAGACCAGGAGGAGAGCAGGTAGAAGGGCTGG 5879
Db 83044 GCCCAGGCCCTCAGCCAGCAGCTCTGTAGACCAGGAGGAGAGCAGGTAGAAGGGCTGG 83103
Qy 5880 CAGGAGTGGAGGTGGAGTGGAGTGGAGAGGAGTCCCTGGGATCTTCTGTGGCCCT 5939
Db 83104 CAGGAGTGGAGGTGGAGTGGAGTGGAGAGGAGTCCCTGGGATCTTCTGTGGCCCT 83163
Qy 5940 TCTGGGTGTCCTCTGTTGGGGCATTTGTGCCAGAGGCACTGGGGTGGCTCCCTGAG 5999
Db 83164 TCTGGGTGTCCTCTGTTGGGGCATTTGTGCCAGAGGCACTGGGGTGGCTCCCTGAG 83223
Qy 6000 GATCCTCTGCCTCACCCTCACTCCAGGCACTTGCCTCAGCCAGACCTACTATCGGGATGA 6059
Db 83224 GATCCTCTGCCTCACCCTCACTCCAGGCACTTGCCTCAGCCAGACCTACTATCGGGATGA 83283
Qy 6060 ACACATCTACCCCTACATGTACCTGGCTGCTACACTGTCCGCAACCGCAATGTGCGGGA 6119
Db 83284 ACACATCTACCCCTACATGTACCTGGCTGCTACACTGTCCGCAACCGCAATGTGCGGGA 83343
Qy 6120 AGCCCTGAGGCCCTGGGGGACAGCGGCACCTGTATCCAGAGGTGAGATCCCCCTACTA 6179
Db 83344 AGCCCTGAGGCCCTGGGGGACAGCGGCACCTGTATCCAGAGGTGAGATCCCCCTACTA 83403
Qy 6180 GGGCTCGAGCCTGCTCTTCTTCTCCCTCCATCATGTTTCCAAACCACTCGTCCAGACT 6239
Db 83404 GGGCTCGAGCCTGCTCTTCTTCTCCCTCCATCATGTTTCCAAACCACTCGTCCAGACT 83463
Qy 6240 GAGCCTGGCTCCGACGCCCATCCCTTTCCATCCAGTCCCTAGGAGCAAGGCCACCA 6299
Db 83464 GAGCCTGGCTCCGACGCCCATCCCTTTCCATCCGCTCCCTAGGAGCAAGGCCACCA 83523
Qy 6300 TTACCCAGGAGGTAGGAGCCCTGATTAAGGTGTACATCTTTCCCTCCCTCCCTCTCCT 6359
Db 83524 TTACCCAGGAGGTAGGAGCCCTGATTAAGGTGTACATCTTTCCCTCCCTCCCTCTCCT 83583
Qy 6360 CCTAATTTTTTTTTTCTCAGAACAGTCTCAAACTCTCAAACTTTTAAACCACCATCATCCAG 6419
Db 83584 CCTAATTTTTTTTTTCTCAGAACAGTCTCAAACTCTCAAACTTTTAAACCACCATCATCCAG 83643
Qy 6420 CAGTGGGACTTCCACCTCGGCCCATCCCTCCCTCTCAATCTTGTCTTCTCTCTCG 6479
Db 83644 CAGTGGGACTTCCACCTCGGCCCATCCCTCCCTCTCAATCTTGTCTTCTCTCTCG 83703
Qy 6480 GCTGACCCAGACAGCATTTTTCAGTGGAGGCCACCTACTTCCCGCAGCCCTCGGG 6539
Db 83704 GCTGACCCAGACAGCATTTTTCAGTGGAGGCCACCTACTTCCCGCAGCCCTCGGG 83763
Qy 6540 GCTCATCCCCCGCAGGTCTCGGGCTACCCCGGTGAGTGTGAGACCCCTTTCAGACCCCTA 6599
Db 83764 GCTCATCCCCCGCAGGTCTCGGGCTACCCCGGTGAGTGTGAGACCCCTTTCAGACCCCTA 83823
Qy 6600 CAGAGACCCCACTGCTCTCAGAGTACAACTACTGCGGGGAGACGAGGAGATCTACAAG 6659
Db 83824 CAGAGACCCCACTGCTCTCAGAGTACAACTACTGCGGGGAGACGAGGAGATCTACAAG 83883
Qy 6660 GAGTCTTTGAGTACCAATGATGTATCCCCAACCTGTGAGGAGGAGGAGGAGCTTG 6719
Db 83984 GAGTCTTTGAGTACCAATGATGTATCCCCAACCTGTGAGGAGGAGGAGGAGCTTG 85022

Db 83984 GAGTCTTTGAAGTAGCAATGATGTATCCCCAACCTGTGAGGAGGAGGAGGAGCTTG 83943
Qy 6720 CTGAGGCGGGCAGGAGCGGCCGGGAGCAAAAGCAGGTGAAAGGCTGAGAGCTCAGC 6779
Db 83944 CTGAGGCGGGCAGGAGCGGCCGGGAGCAAAAGCAGGTGAAAGGCTGAGAGCTCAGC 84003
Qy 6780 CTGFTCCAGCCTCCCACTTGACAGGGCTCCCTTCCACAGGGCCATGGGGCTGCATG 6839
Db 84004 CTGFTCCAGCCTCCCACTTGACAGGGCTCCCTTCCACA -GGCCATGGGGCTGCATGT 84062
Qy 6840 ACGGGATTAGGGATGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGC 6899
Db 84063 ACGGGATTAGGGATGGCAGGAGGAGGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGC 84122
Qy 6900 TATAACTGAGGTCTCTGGGCCACCTTGGACGGGACTGAAGCTATTTTAGAGGTTTCTACC 6959
Db 84123 TATAACTGAGGTCTCTGGGCCACCTTGGACGGGACTGAAGCTATTTTAGAGGTTTCTACC 84182
Qy 6960 CTGTGCTTACAGTTTCAATGGCCAGACTCCCTCCCTCAGCTGAGGGGTGAGGTAGGGATG 7019
Db 84183 CTGTGCTTACAGTTTCAATGGCCAGACTCCCTCCCTCAGCTGAGGGGTGAGGTAGGGATG 84242
Qy 7020 GTACGTCTGCTGCTATGATTTGGCTTTTATAAAGGAGAGAGGTCTTAAGAATGTTCCCAAC 7079
Db 84243 GTACGTCTGCTGCTATGATTTGGCTTTTATAAAGGAGAGAGGTCTTAAGAATGTTCCCAAC 84302
Qy 7080 CTATGCTTACCTTTTCTGGAGCCAGGGCTTTGCTAGGTGGGGGCTGGCCCTGTGCC 7139
Db 84303 CTATGCTTACCTTTTCTGGAGCCAGGGCTTTGCTAGGTGGGGGCTGGCCCTGTGCC 84362
Qy 7140 CTCTGCTAAGGGGTGAGTAAAGAGACTGATCTGTGCTCCCTTCCCTCCCTCCCTCCAGGBCA 7199
Db 84363 CTCTGCTAAGGGGTGAGTAAAGAGACTGATCTGTGCTCCCTTCCCTCCCTCCAGGBCA 84422
Qy 7200 CCCAGAGCCAAAGTTTCCGCCCTCCAGGACCTGAGTGTGCTCCGCCACCTGCTGGGATCT 7259
Db 84423 CCCAGAGCCAAAGTTTCCGCCCTCCAGGACCTGAGTGTGCTCCGCCACCTGCTGGGATCT 84482
Qy 7260 ACAGCGCATCTCGAAATGGGAGGAGGAGTCCACGCCCTGTGCTGACCTGGGCTGGG 7319
Db 84483 ACAGCGCATCTCGAAATGGGAGGAGGAGTCCACGCCCTGTGCTGACCTGGGCTGGG 84542
Qy 7320 CCACCTTTCTGTGCTCAGTCCCTTAGGCCGTTTTGGAGGACAGGTGAGGAGACAGTGCACAG 7379
Db 84543 CCACCTTTCTGTGCTCAGTCCCTTAGGCCGTTTTGGAGGACAGGTGAGGAGACAGTGCACAG 84602
Qy 7380 AGGTCTGGGCACTACAGTGTGACAGCAGCCAGGGCTTGTCTCAGACTTTTCTGCCACAG 7439
Db 84603 AGGTCTGGGCACTACAGTGTGACAGCAGCCAGGGCTTGTCTCAGACTTTTCTGCCACAG 84662
Qy 7440 GGGCAGCATCTGCCCATCCCTTCCGGTCCCGATGGGACTGAGACCCCTGGGTGGGATGG 7499
Db 84663 GGGCAGCATCTGCCCATCCCTTCCGGTCCCGATGGGACTGAGACCCCTGGGTGGGATGG 84722
Qy 7500 GATGGCAGACAGGGTCTTGAGTTCAGCCACTTGGCCGCAACCTTGTCTCAGCTTGC 7559
Db 84723 GATGGCAGACAGGGTCTTGAGTTCAGCCACTTGGCCGCAACCTTGTCTCAGCTTGC 84782
Qy 7560 CTCTCCCCACTGGCCAGGTGCGGCAGAGGTGCGCATAGTGAGCCGAGAGAGCCGAGCGG 7619
Db 84783 CTCTCCCCACTGGCCAGGTGCGGCAGAGGTGCGCATAGTGAGCCGAGAGAGCCGAGCGG 84842
Qy 7620 GCCGAGGCCGAGGAGCCGTGGGGGAGGAGGAGCCCGGAGGAGCCGCGGGGGCCAGCGG 7679
Db 84843 GCCGAGGCCGAGGAGCCGTGGGGGAGGAGGAGCCCGGAGGAGCCGCGGGGGCCAGCGG 84902
Qy 7680 CGGAGGTCCAAAGCAGAGGAGCCCGCCCGCCCAAGAGCCAGCACTGGCAAGGCGCTG 7739
Db 84903 CGGAGGTCCAAAGCAGAGGAGCCCGCCCGCCCAAGAGCCAGCACTGGCAAGGCGCTG 84962
Qy 7740 GGCACCGGCCAGGTGAGTGTGAGGACCCCGCCCGGAGGAGCCCTTCTGGGACTGTGCTGCGC 7799
Db 84963 GGCACCGGCCAGGTGAGTGTGAGGACCCCGCCCGGAGGAGCCCTTCTGGGACTGTGCTGCGC 85022

```
QY 7800 ACAGCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCAGCACCAGCAGCATCACCA 7859
|||||
Db 85023 ACAGCCGAGGCGCTGAAGTGGCAGCAGCGCTCAGGTGCCAGCACCAGCAGCATCACCA 85082
|||||
QY 7860 CGCCCGAGGCTCCAGTGTCTACTTCCAGAGTGAGAGATGAAGGCGCATGAAGAGAGCTG 7919
|||||
Db 85083 CGCCCGAGGCTCCAGTGTCTACTTCCAGAGTGAGAGATGAAGGCGCATGAAGAGAGCTG 85142
|||||
QY 7920 CTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTCAACTCGACGACGACGACGCGCA 7979
|||||
Db 85143 CTGGTGGCCACCAAGATCAACTCGAGCGCCATCAAGCTCAACTCGACGACGACGACGCGCA 85202
|||||
QY 7980 GTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTCTCTCAAG 8039
|||||
Db 85203 GTGCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTACTACACTCTGTCTTCTCTCAAG 85262
|||||
QY 8040 CGGCAGCGCAAGGCGCTCTGAAGTACTGGGAGCTTCGGACCGCTTGTGGGACCCAGGCT 8099
|||||
Db 85263 CGGCAGCGCAAGGCGCTCTGAAGTACTGGGAGCTTCGGACCGCTTGTGGGACCCAGGCT 85322
|||||
QY 8100 CGGCCTTAGTCCCGCAACTCTGAGCCCATGTCTGCCCCCAGCCCAAGGGGACAGGCT 8159
|||||
Db 85323 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 85382
|||||
QY 8160 CACCTCTACCCAAACCTAGTTCGCGTCCCGAGTACAGTCTGTATCAAAACCCAGCTT 8219
|||||
Db 85383 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 85442
|||||
QY 8220 TTCCTCAGCTCAGAACCCAGGCTCTGCCCCAGTCTGTTAGATATAGTCTCTT 8273
|||||
Db 85443 NNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 85496
|||||
```

RESULT 4

```
AP000928/c
LOCUS AP000928 171980 bp DNA HTG 30-MAY-2000
DEFINITION Homo sapiens chromosome 11 clone RP11-772K10 map 11q13, WORKING
DRAFT SEQUENCE, 33 unordered pieces.
ACCESSION AP000928
VERSION AP000928.2 GI:8119069
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 171980)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 171,980 genomic DNA of 11q13
Published Only in Database (1999) In press
2 (bases 1 to 171980)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (17-DEC-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamiara, Kanagawa 228-8555,
Japan [E-mail:hattori@gsr.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924]
On May 31, 2000 this sequence version replaced gi:6997744.
----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsr.riken.go.jp
----- Project Information
Center project name: HumDraft11
Center clone name: Rpl1-772K10
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
```

Assembly program: Phrap; version 0.990329

Consensus quality: 141099 bases at least Q40

Consensus quality: 154862 bases at least Q30

Consensus quality: 163556 bases at least Q20

Insert size: 168780; sum-of-contigs

Quality coverage: 3.98x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known, and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```
1 13598 contig of 13598 bp in length
28815 contig of 13117 bp in length
37862 contig of 10947 bp in length
46388 contig of 8426 bp in length
55060 contig of 8572 bp in length
61498 contig of 6338 bp in length
68961 contig of 7363 bp in length
76353 contig of 7292 bp in length
84390 contig of 7937 bp in length
91104 contig of 6614 bp in length
97347 contig of 6143 bp in length
103810 contig of 6363 bp in length
110422 contig of 6512 bp in length
116801 contig of 6279 bp in length
121989 contig of 5088 bp in length
125946 contig of 3857 bp in length
132109 contig of 6063 bp in length
136422 contig of 4213 bp in length
140566 contig of 4044 bp in length
144574 contig of 3908 bp in length
148311 contig of 3637 bp in length
150652 contig of 2241 bp in length
153335 contig of 2583 bp in length
155668 contig of 2233 bp in length
158851 contig of 3083 bp in length
161241 contig of 2290 bp in length
163307 contig of 1966 bp in length
163408 contig of 2080 bp in length
167734 contig of 2147 bp in length
168855 contig of 1021 bp in length
170089 contig of 1134 bp in length
170955 contig of 466 bp in length
171980 contig of 1225 bp in length
```

Sequence updated (26-May-2000).

* NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

```
1 13598 contig of 13598 bp in length
13599 13598: gap of 100 bp
13699 26815: contig of 13117 bp in length
26816 26915: gap of 100 bp
26916 37862: contig of 10947 bp in length
37863 37962: gap of 100 bp
37963 46388: contig of 8426 bp in length
46389 46488: gap of 100 bp
46489 55060: contig of 8572 bp in length
55061 55160: gap of 100 bp
55161 61498: contig of 6338 bp in length
61499 61598: gap of 100 bp
61599 68961: contig of 7363 bp in length
68962 69061: gap of 100 bp
69062 76353: contig of 7292 bp in length
76354 76453: gap of 100 bp
76454 84390: contig of 7937 bp in length
```

```
* 84391 84490: gap of 100 bp
* 84491 91104: contig of 6614 bp in length
* 91105 91204: gap of 100 bp
* 91205 97347: contig of 6143 bp in length
* 97348 97447: gap of 100 bp
* 97448 103810: contig of 6363 bp in length
* 103811 103910: gap of 100 bp
* 103911 110422: contig of 6512 bp in length
* 110423 110522: gap of 100 bp
* 110523 116801: contig of 6279 bp in length
* 116802 116901: gap of 100 bp
* 116902 121989: contig of 5088 bp in length
* 121990 122089: gap of 100 bp
* 122090 125946: contig of 3857 bp in length
* 125947 126046: gap of 100 bp
* 126047 132109: contig of 6063 bp in length
* 132110 132209: gap of 100 bp
* 132210 136422: contig of 4213 bp in length
* 136423 136522: gap of 100 bp
* 136523 140566: contig of 4044 bp in length
* 140567 140666: gap of 100 bp
* 140667 144574: contig of 3908 bp in length
* 144575 144674: gap of 100 bp
* 144675 148311: contig of 3637 bp in length
* 148312 148411: gap of 100 bp
* 148412 150652: contig of 2241 bp in length
* 150653 150752: gap of 100 bp
* 150753 153335: contig of 2583 bp in length
* 153336 153435: gap of 100 bp
* 153436 155668: contig of 2233 bp in length
* 155669 155768: gap of 100 bp
* 155769 158851: contig of 3083 bp in length
* 158852 158951: gap of 100 bp
* 158952 161241: contig of 2290 bp in length
* 161242 161341: gap of 100 bp
* 161342 163307: contig of 1966 bp in length
* 163308 163407: gap of 100 bp
* 163408 165487: contig of 2080 bp in length
* 165488 165587: gap of 100 bp
* 165588 167734: contig of 2147 bp in length
* 167735 167834: gap of 100 bp
* 167835 168855: contig of 1021 bp in length
* 168856 168955: gap of 100 bp
* 168956 170089: contig of 1134 bp in length
* 170090 170189: gap of 100 bp
* 170190 170655: contig of 466 bp in length
* 170656 170755: gap of 100 bp
* 170756 171980: contig of 1225 bp in length.
```

FEATURES

source

```
1. .171980
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="RP11-772K10"
1. .13598
/feature="assembly_fragment"
13699. .26815
/feature="assembly_fragment"
26916. .37862
/feature="assembly_fragment"
37963. .46388
/feature="assembly_fragment"
46489. .55060
/feature="assembly_fragment"
55161. .61498
/feature="assembly_fragment"
61599. .68961
/feature="assembly_fragment"
69062. .76353
/feature="assembly_fragment"
76454. .84390
/feature="assembly_fragment"
```

```
misc_feature 84491. .91104
/feature="assembly_fragment"
misc_feature 91205. .97347
/feature="assembly_fragment"
misc_feature 97448. .103810
/feature="assembly_fragment"
misc_feature 103911. .110422
/feature="assembly_fragment"
misc_feature 110523. .116801
/feature="assembly_fragment"
misc_feature 116902. .121989
/feature="assembly_fragment"
misc_feature 122090. .125946
/feature="assembly_fragment"
misc_feature 126047. .132109
/feature="assembly_fragment"
misc_feature 132210. .136422
```

Query Match 63.7%; Score 5844; DB 82; Length 171980;
Best Local Similarity 97.7%; Pred. No. 0;
Matches 6036; Conservative 0; Mismatches 120; Indels 20; Caps 16;

```
QY 2421 AGCATTTTCTGGCTGTCACACGGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGC 2480
DB 132109 AGCATTTTCTGGCTGTCACACGGCTCATCCCTACCAACGTTCCCGAGCTCACCTTCCAGC 132050
QY 2481 CCAGCCCCCGCCCGACCGCTGGCGGCTCAGCTACTTTTCCCGTGGCGGACCTGTCTCA 2540
DB 132049 CCAGCCCCCGCCCGACCGCTGGCGGCTCAGCTACTTTTCCCGTGGCGGACCTGTCTCA 131990
QY 2541 TCATCGCGCGCTCTATGCGCGCTTACCGCCAGATCCGAGCGCGCTGACCTGTGCC 2600
DB 131989 TCATCGCGCGCTCTATGCGCGCTTACCGCCAGATCCGAGCGCGCTGACCTGTGCC 131930
QY 2601 TCTATCTCTGAGAGGGGGTGTCTCCAGCGCTGAGCTGGTGAAGAGGTCTCCCATGTCA 2660
DB 131929 TCTATCTCTGAGAGGGGGTGTCTCCAGCGCTGAGCTGGTGAAGAGGTCTCCCATGTCA 131870
QY 2661 TATGGAACAGCTCAGCGCTCTCTACTTCAAGAGTCCGGCCACATCCAGTCCCTCTTCA 2720
DB 131869 TATGGAACAGCTCAGCGCTCTCTACTTCAAGAGTCCGGCCACATCCAGTCCCTCTTCA 131810
QY 2721 GCTTCATCAGAGTTGGAGCCAGTAGTGGGAATCTTATCCATGACCCAGCTTCTTCAAA 2780
DB 131809 GCTTCATCAGAGTTGGAGCCAGTAGTGGGAATCTTATCCATGACCCAGCTTCTTCAAA 131750
QY 2781 ACCCTCCATGGTTTACAGACCCCTTTTAAAGAACTGTAAGCCTTGTGAGGTTTGGCAGGTG 2840
DB 131749 ACCCTCCATGGTTTACAGACCCCTTTTAAAGAACTGTAAGCCTTGTGAGGTTTGGCAGGTG 131690
QY 2841 TTATTTTCTCTTTTGCAGTTGGGAACTGAAAGCCAGAGGGGAAATGATATGCCAAAG 2900
DB 131689 TTATTTTCTCTTTTGCAGTTGGGAACTGAAAGCCAGAGGGGAAATGATATGCCAAAG 131630
QY 2901 TCACACAGGGCATGGCAGGGCTGGAAGTGAAGCCTGATCATCTTGGCTCCCAATCATCAAC 2960
DB 131629 TCACACAGGGCATGGCAGGGCTGGAAGTGAAGCCTGATCATCTTGGCTCCCAATCATCAAC 131570
QY 2961 CTCACCTTCTGCCCTCTCAGCACCCCTTGGCAGTGAACAGCTACAGGAGTTCTTAAG 3020
DB 131569 CTCACCTTCTGCCCTCTCAGCACCCCTTGGCAGTGAACAGCTACAGGAGTTCTTAAG 131510
QY 3021 CATGAGACACAGGGGGCGGACGAGATTTAGGGGGCAAGAGATCAAAATGGGCTTGCATT 3080
DB 131509 CATGAGACACAGGGGGCGGACGAGATTTAGGGGGCAAGAGATCAAAATGGGCTTGCATT 131450
QY 3081 TGAGGCAGTTAAACAAATAATGCTATGAAGATTTTTTTTTTTTTTTTTTTTTTTTGGAGACA 3140
DB 131449 TGAGGCAGTTAAACAAATAATGCTATGAAGA-TTTTTTTTTTTTTTTTTTTTTTTTGGAGACA 131391
QY 3141 GGGTCTCAGTCTGTCCCGGCTGGAGTGCAGTGGTGTGATCATGGCTCACTCAGCAGCT 3200
DB 131390 GGGTCTCAGTCTGTCCCGGCTGGAGTGCAGTGGTGTGATCATGGCTCACTCAGCAGCT 131331
```

QY 3201 CAGTCTCCCTGGGCTCAGAGATCCTCCAAACCTCAGCCTCCTGAGTAGCTGAGAGTACAGG 3260
|||||
Db 131330 CAGTCTCCCTGGGCTCAGAGATCCTCCAAACCTCAGCCTCCTGAGTAGCTGAGAGTACAGG 131271

QY 3261 CAGTCACCGTGGTGGTAAATTTTTGTATTTTTTTTGTAGAGATGAGTCTCACTAT 3320
|||||
Db 131270 CATGCACCGTGGTGGTAAATTTTTTTGTATTTTTTTGTAGAGATGAGTCTCACTAT 131211

QY 3321 GTGGCCACAGCTGGCTCTCAACTCTTTGGCTCAAGTGATCTGCCGCCCTCAGTCTCCCAA 3380
|||||
Db 131210 GTGGCCACAGCTGGCTCTCAACTCTTTGGCTCAAGTGATCTGCCGCCCTCAGTCTCCCAA 131151

QY 3381 ATGCTGGGATTTACAGGTGTGAGCCACCGCAACTGTGGCTATGAAA-ATTTTTTTTTTT 3439
|||||
Db 131150 ATGCTGGGTTACACGGTGTGAGCCACCGCAACTGTGGCTATGAAAATTTTTTTTTTTT 131091

QY 3440 TTTCACAGCGGTCTCACTCTGTGCGCCAGGCTGGAGTGCAGTGGTGCATACTCGGCTCA 3499
|||||
Db 131090 TTTCACAGCGGTCTCACTCTGTGCGCCAGGCTGGAGTGCAGTGGTGCATACTCGGCTCA 131031

QY 3500 CTGCAAGCTCTGCCTCCTGCTTTTCATGCCATCTCCTGCTCCTGCTCAGCTCCTGAG 3559
|||||
Db 131030 CTGCAAGCTCTGCCTCCTGCTTTTCATGCCATCTCCTGCTCCTGCTCAGCTCCTGAG 130971

QY 3560 TAGCTGGGACTACAGGAGCTGCCACCATGCTGCTAAATTTTTTTTGGATTTTTAGTA 3619
|||||
Db 130970 TAGCTGGGACTACAGGAGCTGCCACCATGCTGCTAAATTTTTTTTGGATTTTTAGTA 130911

QY 3620 GAGACAGAGTTTACCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCGCC 3679
|||||
Db 130910 GAGACAGAGTTTACCATGTTAGCCAGGATGCTCGATCTCCTGACCTCGTGATCCGCC 130851

QY 3680 CGCCTTGGCCTCCCAAGTCTGGGATTACAGCGTGAGCCACCGCAGCTGTCAAAAT 3739
|||||
Db 130850 CGCCTTGGCCTCCCAAGTCTGGGATTACAGCGTGAGCCACCGTACCTGGTCAAAAAT 130791

QY 3740 GTTTGAGACAGAGAGGGCTTTGACCTCAAAAGGCTTAAGAGTCAAGGCTTGCAAAAGAGC 3799
|||||
Db 130790 GTTTGAGACAGAGAGGGCTTTGACCTCAAAAGGCTTAAGAGTCAAGGCTTGCAAAAGAGC 130731

QY 3800 TTTCACCAAGCCCGTTGACTGGCAATCCCATCTGCTGGTGGCCATATTGAGAAGGAATC 3859
|||||
Db 130730 TTTCACCAAGCCCGTTGACTGGCAATCCCATCTGCTGGTGGCCATATTGAGAAGGAATC 130671

QY 3860 AGAGGCTGCTTCTACGTTAGCAGGAAAAGAGTGCAGAGATAAATCAGAGGTTATTTGTG 3919
|||||
Db 130670 AGAGGCTGCTTCTACGTTAGCAGGAAAAGAGTGCAGAGATAAATCAGAGGTTATTTGTG 130611

QY 3920 GTGGGTGTATAGCCAGAGAGTGTGGCCAGCGTCTCTGTTTTTGGCCATTCCTGTTTTAACC 3979
|||||
Db 130610 GTGGGTGTATAGCCAGAGAGTGTGGCCAGCGTCTCTGTTTTTGGCCATTCCTGTTTTAACC 130551

QY 3980 TAGTAGTCAGTAAATGGAATCCCTAAATCCATAGATATATAATAGAGTTGCGAGAGA 4039
|||||
Db 130550 TAGTAGTCAGTAAATGGAATCCCTAAATCCATAGATATATAATAGAGTTGCGAGAGA 130491

QY 4040 AAGACAGCTAGGCCCAAGGCTGGGTGAGCTACAGAGATATCCAGAAAAGTATCTTTGTTG 4099
|||||
Db 130490 AAGACAGCTAGGCCCAAGGCTGGGTGAGCTACAGAGATATCCAGAAAAGTATCTTTGTTG 130431

QY 4100 GACATAGAGGTTAAACAGGAGAGAGTCTTTTGAACAGCTGGAGGGGAAGGATGGAGG 4159
|||||
Db 130430 GACATAGAGGTTAAACAGGAGAGAGTCTTTTGAACAGCTGGAGGGGAAGGATGGAGG 130371

QY 4160 GATAGTGGCAGGAGAAATCTGAGTTGGGTACAGAGCTTTGGAAGGAGTGGGAGGAGT 4219
|||||
Db 130370 GATAGTGGCAGGAGAAATCTGAGTTGGGTACAGAGCTTTGGAAGGAGTGGGAGGAGT 130311

QY 4220 GTGGCCCATCACTACTCG-CCCCTTTCCCATGTTAAAGCAGAGAGGCCCTCTTTTCAT 4278
|||||
Db 130310 GTGGCCCATCACTACTCGCCCCCTTTTCCCATGTTAAAGCAGAGAGGCCCTCTTTTCAT 130251

QY 4279 TACCTCCCTCTTCCACAGGCACCAAAATTTGACAGCTCCGGTGTGGCTTTTTCCTGTGTTG 4338
|||||
Db 130250 TACCTCCCTCTTCCACAGGCACCAAAATTTGACAGCTCCGGTGTGGCTTTGCTGTGGTTG 130191

QY 4339 GGGCTTGCAGAGCCCTTGGGTCTCCGGGATGTCCACCTCGCCCTGTCTGAGGATCATGCCT 4398
|||||
Db 130190 GGGCTTGCAGAGCCCTTGGGTCTCCGGGATGTCCACCTCGCCCTGTCTGAGGATCATGCCT 130131

QY 4399 GGGTAGTGTTTGGGCCCAATGGGGAGCAGACAGCTCAGGTCACTGGCAGCGCAAGGGCA 4458
|||||
Db 130130 GGGTAGTGTTTGGGCCCAATGGGGAGCAGACAGCTCAGGTCACTGGCAGCGCAAGGGCA 130071

QY 4459 ACAGAGACCGCAGGGCCAGACAGTCAATGCCGCTGTGGCTGAGCGGGTATTTTCCCTC 4518
|||||
Db 130070 ACAGAGACCGCAGGGCCAGACAGTCAATGCCGCTGTGGCTGAGCGGGTATTTTCCCTC 130011

QY 4519 CCGCCAGCCTTGTCCCTTTCACTACTGTAGTCCCAAGCCACCAAGGACTCCATTTTC 4578
|||||
Db 130010 CCGCCAGCCTTGTCCCTTTCACTACTGTAGTCCCAAGCCACCAAGGACTCCATTTTC 129951

QY 4579 TTGGGCCACACCCCTTTTCCCATCACCACCCACATAGGAAGGAAGACAGAGAGCCC 4638
|||||
Db 129950 TTGGGCCACACCCCTTTTCCCATCACCACCCACATAGGAAGGAAGACAGAGAGCCC 129891

QY 4639 CTTTTCCTGGCTGTCTATTCCTGAAGCAGGCACAGGCTGGGCCATCATGAGACATAATGA 4698
|||||
Db 129890 CTTTTCCTGGCTGTCTATTCCTGAAGCAGGCACAGGCTGGGCCATCATGAGACATAATGA 129831

QY 4699 TCTCATCCCCCTTAAGAGTGGCTCTACTGAAAGGATCATACATGGCGTGTGACCGCA 4758
|||||
Db 129830 TCTCATCCCCCTTAAGAGTGGCTCTACTGAAAGGATCATACATGGCGTGTGACCGCA 129771

QY 4759 AGATGAGAGTGGGTTCATGGTGTGCCATCAACCTTTCCATTGACCTGCACACCGACT 4818
|||||
Db 129770 AGATGAGAGTGGGTTCATGGTGTGCCATCAACCTTTCCATTGACCTGCACACCGACT 129711

QY 4819 CGCTGAGGCTTCTGACGCTGCAGCAGTGCAGGCTGAGGCTGAGCAATGGGCGAGGACTGGGCTA 4878
|||||
Db 129710 CGCTGAGGCTTCTGACGCTGCAGCAGTGCAGGCTGAGGCTGAGCAATGGGCGAGGACTGGGCTA 129651

QY 4879 GGCCACAGCTTGACTTGTCTGGGACCCCTGGGAGGGGACACTTTCCTTCTGAGCTTCAG 4938
|||||
Db 129650 GGCCACAGCTTGACTTGTCTGGGACCCCTGGGAGGGGACACTTTCCTTCTGAGCTTCAG 129591

QY 4939 CTTCCCTCTCTGAAAATAGGTTAGTAAATCTGCTGGCTTCCCAAGGCTCTTGG 4998
|||||
Db 129590 CTTCCCTCTCTGAAAATAGGTTAGTAAATCTGCTGGCTTCCCAAGGCTCTTGG 129531

QY 4999 GAGAGTAGAATTGAGATGTGAAATTCCTTGACTCCATTAAAGGGCTGGTCCCAAGATTT 5058
|||||
Db 129530 GAGAGTAGAATTGAGATGTGAAATTCCTTGACTCCATTAAAGGGCTGGTCCCAAGATTT 129471

QY 5059 TGGCCCTTCCATGTTGGTGTCTTGGTTCTGACCCCCACCTCTGCCCCGATAGG 5118
|||||
Db 129470 TGGCCCTTCCATGTTGGTGTCTTGGTTCTGACCCCCACCTCTGCCCCGATAGG 129411

QY 5119 CTAAGACCCGTTCTCTCCCTGCTCCGCTCAATAACTCTCTCTCGGCTCCCTAGAA 5178
|||||
Db 129410 CTAAGACCCGTTCTCTCCCTGCTCCGCTCAATAACTCTCTCTCGGCTCCCTAGAA 129351

QY 5179 GCTGCTCTGGCTCTCTATGACCTGGGACATCTGGAAGGTCTAGTAGAGGAAGTGGCCA 5238
|||||
Db 129350 GCTGCTCTGGCTCTCTATGACCTGGACATCTGGAAGGTCTAGTAGAGGAAGTGGCCA 129291

QY 5239 GGTGCGCCTGTGAGCGCGGGGCTGGGTGGCAGCCTGAATATGATCTCTTTCCCTAGG 5298
|||||
Db 129290 GGTGCGCCTGTGAGCGCGGGGCTGGGTGGCAGCCTGAATATGATCTCTTTCCCTAGG 129231

QY 5299 TACCCCATGSCCTTAGGAACTGGCAGATCTAGAGAGCTGAGCCCAACCCCTGGCCGG 5358
|||||
Db 129230 TACCCCATGSCCTTAGGAACTGGCAGATCTAGAGAGCTGAGCCCAACCCCTGGCCGG 129171

QY 5359 CCAGACCCACTCACCCCTCTACCAACAAGGTGGGGCTACTAAGGAGGGTGCAGAGGGAGA 5418

AP000575
AP000575.3 GI:8118783
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens DNA, clone:XXpl-95F6.

1 (bases 1 to 73431)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T. D., Hong-Seong, P.

Published Only in DataBase (1999) In press
2 (bases 1 to 73431)

Submitted (08-OCT-1999) Masahira Hattori The Institute of Physical and Chemical Research

Japan (E-mail: hattori@gsc.riken.go.jp,
<http://hsc.riken.go.jp/~hatori>, Tel: 01 43 770 0022)

On May 31, 2000 this sequence version replaced [g1.0557401](#).
----- Genome Center

web site: <http://nyp.gsc.riken.go.jp/>
Contact: hattori@gsc.riken.go.jp

Center clone name: xyp1-95f6
----- Summary Statistics

Assembly program: Phrap; version 0.990329
Consensus quality: 66753 bases at least Q40

Insert size: 7031; sum-of-contigs
Quality coverage: 4.73x in Q20 bases; sum-of-contigs

15 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the

as soon as it is available and the accession number will be preserved

Accession	Contig	Length (bp)
20768	29453	8686
29554	37145	7592

47445	52652 contig of	5208 bp in length
52753	56084 contig of	3332 bp in length

Accession	Contig	Length (bp)
63600	66501	2902
66602	68545	1944

72389 73431 contig of 1043 bp in length
Sequence updated (12-Oct-1999)

- * consists of 15 contigs. The true order of the pieces
- * is not known and their order in this sequence record is

* This record will be updated with the finished sequence

Db	17886	GGGCACGAGACCGCAGGGGCCACAGTCAATGCCGGTGTGGCTGAGCGGGTATTGT	1782
Qy	4514	CCCTCCCCCAGCCTTGTCCTTCATACTGTAGTACCCCAAGCCACCAAGGACTCCA	4573
Db	17826	CCCTCCCCCAGCCTTGTCCTTCATACTGTAGTACCCCAAGCCACCAAGGACTCCA	17767
Qy	4574	TTTTCTTGGGCCACACCCCTTTCTTCCATCACACCCACATAGAAAGGACAGAA	4633
Db	17766	TTTTCTTGGGCCACACCCCTTTCTTCCATCACACCCACATAGAAAGGACAGAA	17707
Qy	4634	AGCCCCCTTTCCTGGCTGTCTAATCCCTGAAGCAGGCACAGGTTGGCCATCATGAGACAT	4693
Db	17706	AGCCCCCTTTCCTGGCTGTCTAATCCCTGAAGCAGGCACAGGTTGGCCATCATGAGACAT	17647
Qy	4694	AATGATCTCATCCCCCTTAAGAGCTGGCTACTCTGAAGAGTATACATGCGCTGTGA	4753
Db	17646	AATGATCTCATCCCCCTTAAGAGCTGGCTACTCTGAAGAGTATACATGCGCTGTGA	17587
Qy	4754	CCGCAAGATGAGGTGGGGTTTCATGGTGTGCCATCAACCTTCCATPTGACCTGCAC	4813
Db	17586	CCGCAAGATGAGGTGGGGTTTCATGGTGTGCCATCAACCTTCCATPTGACCTGCAC	17527
Qy	4814	CGACTCGCTGAGCTTCTGCAAGCTGCACAGGTGAGGCTCAGCCAATGGGCGAGGACTG	4873
Db	17526	CGACTCGCTGAGCTTCTGCAAGCTGCACAGGTGAGGCTCAGCCAATGGGCGAGGACTG	17467
Qy	4874	GGCTAGGCCACACTTGACTTCTGTGGGACCTTGGGCAAGGGCACTTTCCCTTCTGTAGC	4933
Db	17466	GGCTAGGCCACACTTGACTTCTGTGGGACCTTGGGCAAGGGCACTTTCCCTTCTGTAGC	17407
Qy	4934	TTGAGCTTCCCTTCTTGGAAAAATGGGTAGTAATTCCTGGCCCTGGCCTTTCCAGGSGCT	4993
Db	17406	TTGAGCTTCCCTTCTTGGAAAAATGGGTAGTAATTCCTGGCCCTGGCCTTTCCAGGSGCT	17347
Qy	4994	CTTGGGAGAGTAGAATTGAGATGTCAATTCCTTTGACTCCATTAAGGGGCTGGTCCAG	5053
Db	17346	CTTGGGAGAGTAGAATTGAGATGTCAATTCCTTTGACTCCATTAAGGGGCTGGTCCAG	17287
Qy	5054	AATTTTGGCCCTTCCACATGCTGGGTGGTCCCTGTGGTTCCTGACCCCCACCTCTGCCCCG	5113
Db	17286	AATTTTGGCCCTTCCACATGCTGGGTGGTCCCTGTGGTTCCTGACCCCCACCTCTGCCCCG	17227
Qy	5114	ATAGCTTAAGACCCGTTCTCCTCCCTGTTCGGTGGCTCATACCTCTCCTTCGGCTCC	5173
Db	17226	ATAGCTTAAGACCCGTTCTCCTCCCTGTTCGGTGGCTCATACCTCTCCTTCGGCTCC	17167
Qy	5174	TAGAACTGTCTTGGCTGCTCTATGACCTTGGACATCTGGAAGGTCACTAGAGGGAAGT	5233
Db	17166	TAGAACTGTCTTGGCTGCTCTATGACCTTGGACATCTGGAAGGTCACTAGAGGGAAGT	17107
Qy	5234	GGCAGGCTGGCTGTGGTGGCCGGGGGCTGGGTGGCAGCCTGAATTATGATCTTTTC	5293
Db	17106	GGCAGGCTGGCTGTGGTGGCCGGGGGCTGGGTGGCAGCCTGAATTATGATCTTTTC	17047
Qy	5294	CTAGGTACCCCATGGCCCTTAGGGAACCTTGGCAGATCTTAGAGAGCTGGAGCCCCACCCCTG	5353
Db	17046	CTAGGTACCCCATGGCCCTTAGGGAACCTTGGCAGATCTTAGAGAGCTGGAGCCCCACCCCTG	16987
Qy	5354	GCCGCCAGACCCACTCACCTCTACACACAGGTGGGGGCATCTTAAGAGGGTGCAGAA	5413
Db	16986	GCCGCCAGACCCACTCACCTCTACACACAGGTGGGGGCATCTTAAGAGGGTGCAGAA	16927
Qy	5414	GGAGACCTTAACAGTGGCTGAGGACGGGGCCCTCATCTGGGCAGATGAGAAGAACATTT	5473
Db	16926	GGAGACCTTAACAGTGGCTGAGGACGGGGCCCTCATCTGGGCAGATGAGAAGAACATTT	16867
Qy	5474	GTGTGTTGGGGGATCGCCCATCCAGTCTCACATTTGTGTCAACTGTGTGCAAGATCACT	5533
Db	16866	GTGTGTTGGGGGATCGCCCATCCAGTCTCACATTTGTGTCAACTGTGTGCAAGATCACT	16807
Qy	5534	TCAGTCAAGGCTGTCTGAGGGGTGTCCAGGTTCCCCAGCTGGGAGTGGCAGGGGGCTGC	5593
Db	16806	TCAGTCAAGGCTGTCTGAGGGGTGTCCAGGTTCCCCAGCTGGGAGTGGCAGGGGGCTGC	16747

QY	5594	ATTTCGTCCTCCCTCAGCCCTGCCTTTTCTTCGCCACTGCTTACTGCTCCTTCCTGGAGTATAACA	5653
DB	16746	ATTTCGTCCTCCCTCAGCCCTGCCTTTTCTTCGCCACTGCTTACTGCTCCTTCCTGGAGTATAACA	16687
QY	5654	GAGGTCAAATGTGGTFAAGAGCACGTGAAGAGGGGGTTTCACTTGGTGGGTGTAGCTGGGG	5713
DB	16686	GAGGTCAAATGTGGTFAAGAGCACGTGAAGAGGGGGTTTCACTTGGTGGGTGTAGCTGGGG	16627
QY	5714	AGGAGGGCCATTGGGCTGGGCTTGAAGTCTTTTGGTGTATGTGTAGAAGAGTGTCTTGAGAA	5773
DB	16626	AGGAGGGCCATTGGGCTGGGCTTGAAGTCTTTTGGTGTATGTGTAGAAGAGTGTCTTGAGAA	16567
QY	5774	AGAGAAGGGCCCTGAGCTCGAGGGCAGGCCACCCCTGCAGTCTGCCCAGGCCTCAG	5833
DB	16566	AGAGAAGGGCCCTGAGCTCGAGGGCAGGCCACCCCTGCAGTCTGCCCAGGCCTCAG	16507
QY	5834	CCAGCAGTCCCTGTAGACCCACGGGAGGACACGAGTAGAAGGGCTGGCAGCAGTGGAGGT	5893
DB	16506	CCAGCAGTCCCTGTAGACCCACGGGAGGACACGAGTAGAAGGGCTGGCAGCAGTGGAGGT	16447
QY	5894	GGGAGTGAAGATGGAGAGACTCCCTGGGATCTTTCCTGTGGCCCTTCTGGGTGTGCCCT	5953
DB	16446	GGGAGTGAAGATGGAGAGACTCCCTGGGATCTTTCCTGTGGCCCTTCTGGGTGTGCCCT	16387
QY	5954	GGTGGGGCATTTGTCCACAGCAGGGCAGCTGGGGCTGCCTCTGAGGATCTCTTCGCCCTCA	6013
DB	16386	GGTGGGGCATTTGTCCACAGCAGGGCAGCTGGGGCTGCCTCTGAGGATCTCTTCGCCCTCA	16327
QY	6014	CCTCCATCCAGGGCATTCGCTCAGCCAAAGACTACTATCGGGATGAACACATCTACCCCT	6073
DB	16326	CCTCCATCCAGGGCATTCGCTCAGCCAAAGACTACTATCGGGATGAACACATCTACCCCT	16267
QY	6074	ACATCTACCTGCCTGGCTACCACTGTCCCAACCGCAATGTGGGGAAGCCCTGCAGGCT	6133
DB	16266	ACATCTACCTGCCTGGCTACCACTGTCCCAACCGCAATGTGGGGAAGCCCTGCAGGCT	16207
QY	6134	GGCGGCACAGGCCACTGTCATCCAGGAGTGAGGATCCCCCTACTTAGGGCCTGCAGCCTG	6193
DB	16206	GGCGGCACAGGCCACTGTCATCCAGGAGTGAGGATCCCCCTACTTAGGGCCTGCAGCCTG	16147
QY	6194	TCCCTTCTTCCCTCCATCAGTTTCCACACAGCCCTGTCAGGACTGAGGCTGCCTCCC	6253
DB	16146	TCCCTTCTTCCCTCCATCAGTTTCCACACAGCCCTGTCAGGACTGAGGCTGCCTCCC	16087
QY	6254	ACGCCCATCCCTTTTCCATCCAGTCCCTTAGCAGCAAGGCCACCATTTACCCAGGAGTA	6313
DB	16086	ACGCCCATCCCTTTTCCATCCAGTCCCTTAGCAGCAAGGCCACCATTTACCCAGGAGTA	16027
QY	6314	GGGACCTTGATTAAGGTGTACATCTTTCCTTCCCTCCCTCTCTCTCTTAATTTTTTTTT	6373
DB	16026	GGGACCTTGATTAAGGTGTACATCTTTCCTTCCCTCCCTCCCTCTCTCTTAATTTTTTTTT	15967
QY	6374	TCTCAGAACAGTCTCAATCTCCAATGTTTAAACCACCATCATCCAGCAGTGGGACTTCCA	6433
DB	15966	TCTCAGAACAGTCTCAATCTCCAATGTTTAAACCACCATCATCCAGCAGTGGGACTTCCA	15907
QY	6434	CCCTCGGGCCCATGCCCCCTCTCTCATTTCTTGTCTTCTTGGGCTGACCCAGACAG	6493
DB	15906	CCCTCGGGCCCATGCCCCCTCTCTCATTTCTTGTCTTCTTGGGCTGACCCAGACAG	15847
QY	6494	CATCATTTTGCAGTGAAGACCCCACTACTCCCCCAGCCCCC-TGGGGGCTCCATCCCCCG	6552
DB	15846	CATCATTTTGCAGTGAAGACCCCACTACTCCCCCAGCCCCGTGGGGCTCCATCCCCCG	15787
QY	6553	CCAGGTCCCTGGGGCTAGCCCCGATGGGTGAGACCCCTTTCAGACCCCTACAGAGACCCCACT	6612
DB	15786	CCAGGTCCCTGGGGCTAGCCCCGATGGGTGAGACCCCTTTCAGACCCCTACAGAGACCCCACT	15727
QY	6613	GCTCTCAGCTACAACCTACTGCGGGGAAGACGAGGAGATCTACAAGGAGTCTTTTGAAG	6672
DB	15726	GCTCTCAGCTACAACCTACTGCGGGGAAGACGAGGAGATCTACAAGGAGTCTTTTGAAG	15667

Qy	6673	TAGCCAATGATGTCATCCCAACCTGCTGAAGAGGAGCAGCCAGCTTGCTGTGAGGCGGGCG	6732
Db	15666	TAGCCAATGATGTCATCCCAACCTGCTGAAGAGGAGCAGCCAGCTTGCTGTGAGGCGGGCG	15607
Qy	6733	AGGAGCGCGCGGGGAGCAAGCCAGCTGAAAGGCTGGAAGGCTCCAGCTCTGTGTCCACGCT	6792
Db	15606	AGGAGCGCGCGGGGAGCAAGCCAGCTGAAAGGCTGGAAGGCTCCAGCTCTGTGTCCACGCT	15547
Qy	6793	CCACCTTGGACAGGGCTCCCTTCCACAGGGCCATGGGGGCTGCACTGACGGGATTAGGGA	6852
Db	15546	CCACCTTGGACAGGGCTCCCTTCCACA - GGCCATGGGGGCTGCATGTACGGGATTAGGGA	15488
Qy	6853	TGGCAGGAGGAAGGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGTATAACTAGAGTC	6912
Db	15487	TGGCAGGAGGAAGGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGTATAACTAGAGTC	15428
Qy	6913	CTGGGCCCCACGTTGGACGGGACTGAAGGTATTTTAGAGGTTTCTACCCCTGTGCCCTTCACT	6972
Db	15427	CTGGGCCCCACGTTGGACGGGACTGAAGGTATTTTAGAGGTTTCTACCCCTGTGCCCTTCACT	15368
Qy	6973	TTTCATGGCCACAGCTCCCTCCCTCAGCTAGGGGTGGAGGTAGGATGGTACGTCCTGGCT	7032
Db	15367	TTTCATGGCCACAGCTCCCTCCCTCAGCTAGGGGTGGAGGTAGGATGGTACGTCCTGGCT	15308
Qy	7033	ATGGATTGGCTTTATAAAGGAAAGAGGTTCTAAGAAATGTTCCCAACCTATGCTTACCTT	7092
Db	15307	ATGGATTGGCTTTATAAAGGAAAGAGGTTCTAAGAAATGTTCCCAACCTATGCTTACCTT	15248
Qy	7093	TTCTGGAGCAGAGGGTCTTTGCTTAGTGGGGGCCCTGGCCCTGTGCCCTCTGTCTAAGGGG	7152
Db	15247	TTCTTGGAGCCAGGGGTCTTTGCCTTAGTGGGGGCCCTGGCCCTGTCTGTCTAAGGGG	15188
Qy	7153	TGAGTAAGAGACTGATCTGTCCCTCCCTTCCCTCCCTCGTCCAGGGCACCCAGAGCCCAAGG	7212
Db	15187	TGAGTAAGAGACTGATCTGTCCCTCCCTTCCCTCCCTCGTCCAGGGCACCCAGAGCCCAAGG	15128
Qy	7213	TTCCGCCCTCCAGGACCTCAGTCTTCGCCACCTGCTGCGATTCTACGACGGCATCTG	7272
Db	15127	TTCCGCCCTCCAGGACCTCAGTCTTCGCCACCTGCTGCGATTCTACGACGGCATCTG	15068
Qy	7273	CAAAATGGAGAGGGCAGTCCACGCTGTGCTGCAGCTGGGCTGGGCCACCTTCTTTGT	7332
Db	15067	CAAAATGGAGAGGGCAGTCCACGCTGTGCTGCAGCTGGGCTGGGCCACCTTCTTTGT	15008
Qy	7333	GCAGTCCCTAGGCCGTTTTGAGGGACAGGTGAGGGACAGCTGCACAGAGGTCTCGGGCACT	7392
Db	15007	GCAGTCCCTAGGCCGTTTTGAGGGACAGGTGAGGGACAGCTGCACAGAGGTCTCGGGCACT	14948
Qy	7393	ACAGGTGGTGACAGCAGCCACGGGCTTGTACAGATTTTCTGGCCAGGGGCGAGCATCTGC	7452
Db	14947	ACAGGTGGTGACAGCAGCCACGGGCTTGTACAGATTTTCTGGCCAGGGGCGAGCATCTGC	14888
Qy	7453	CCATCCCTTCGGTGGCGATGGGACTGAGACCCCTGGTGGGATGGATGGCCAGAGCA	7512
Db	14887	CCATCCCTTCGGTGGCGATGGGACTGAGACCCCTGGTGGGATGGATGGCCAGAGCA	14828
Qy	7513	GGGTCTGGAGTTCACGCCACTGGCCGCACTTGTCTCACCTGTCTCCTCCCACTGG	7572
Db	14827	GGGTCTGGAGTTCACGCCACTGGCCGCACTTGTCTCACCTGTCTCCTCCCACTGG	14768
Qy	7573	CCCAGGTGCGGCAGAAAGTGGCATAGTGAAGCCAGAGCCGAGCGCGGCCAGCCGAGG	7632
Db	14767	CCCAGGTGCGGCAGAAAGTGGCATAGTGAAGCCAGAGCCGAGCGCGGCCAGCCGAGG	14708
Qy	7633	AGCCGTGGGGGAGGAAGCCCGGAAGCCCGGCGGGGGCCCCACGGGGGAGTCTCAAGC	7692
Db	14707	AGCCGTGGGGGAGGAAGCCCGGAAGCCCGGCGGGGGCCCCACGGGGGAGTCTCAAGC	14648
Qy	7693	CAGAGGACCCCGCCGCCCAAGAGCAGCTGGAACAGGGCTTGGCACCCGCCAGG	7752
Db	14647	CAGAGGACCCCGCCGCCCAAGAGCAGCTGGAACAGGGCTTGGCACCCGCCAGG	14588
Qy	7753	GTGCAGTGTGACGACACCCCGGAAGCTCTCGTGGAGTGTCTGTGGACAGCCCGAGGCC	7812

Db	14587	GTGAGTGTCTCAGGACCCCCCGGAAGCCCTCTCTGGACTCTGCGTGGCACAGCCCGCAGGCC	14598
Qy	7813	CTGAAGGTGGCAGCAGCGCTCAGGTGCGCAGCACCGCAGCATCACACCGCGGAGGGTC	7872
Db	14527	CTGAAGGTGGCAGCAGCGCTCAGGTGCGCAGCACCGCAGCATCACACCGCGGAGGGTC	14468
Qy	7873	CAGTGTCTCACTTTCCAGAGTGAAGAGTGAAGGGCATGAAGGAGTGTGTGGTGGCCACCA	7932
Db	14467	CAGTGTCTCACTTTCCAGAGTGAAGAGTGAAGGGCATGAAGGAGTGTGTGGTGGCCACCA	14408
Qy	7933	AGATCAACTCGAGCGCCATCAAGCTGCAACTTCAGGCACAGTCGCAAGTCGAGATGAAGA	7992
Db	14407	AGATCAACTCGAGCGCCATCAAGCTGCAACTTCAGGCACAGTCGCAAGTCGAGATGAAGA	14348
Qy	7993	AGCAGAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTTCTCTCAAGCGGAGCGCAAG	8052
Db	14347	AGCAGAAAGTGTCCACCCCTAGTGACTACACTCTGTCTTTCTCTCAAGCGGAGCGCAAG	14288
Qy	8053	GCCTCTGAACACTCTGGGACCTTCGGACCGCTTGTGGGACCGCAGGCTCCG-CCTTAGTCC	8111
Db	14287	GCCTCTGAACACTCTGGGACCTTCGGACCGCTTGTGGGACCGCAGGCTCCGCGCTTAGTCC	14228
Qy	8112	CCCAACTCTGAGCCCATGTTCTGCCCCCAGCCCCAAAGGGGACAGGCCTCTCTACCCA	8171
Db	14227	CCCAACTCTGAGCCCATGTTCTGCCCCCAGCCCCAAAGGGGACAGGCCTCTCTACCCA	14168
Qy	8172	AACCTTAGTTCCTCCGCTCCGAGTACAGTCTGTATCAAAACCCAGCATTTTCTCCAGCTCA	8231
Db	14167	AACCTTAGTTCCTCCGCTCCGAGTACAGTCTGTATCAAAACCCAGCATTTTCTCCAGCTCA	14108
Qy	8232	GAACCCAGGGCTCTGCCCCAGTCGGTTAGAATATAGGTCTCTCTTCCCAAGAAATCCACGCG	8291
Db	14107	GAACCCAGGGCTCTGCCCCAGTCGGTTAGAATATAGGTCTCTCTTCCCAAGAAATCCACGCG	14048
Qy	8292	GCCAATGGAAACCTCACGCTGGGTCTTAATTACCAAGTCTTTTAAAGGCCACGCCCTAGAA	8351
Db	14047	GCCAATGGAAACCTCACGCTGGGTCTTAATTACCAAGTCTTTTAAAGGCCACGCCCTAGAA	13988
Qy	8352	ACCCAGCTCTCTCTCGAAGCGCTCACTTAGAGCCAGCACCAAGCTTACTCAGGGCTCCT	8411
Db	13987	ACCCAGCTCTCTCTCGAAGCGCTCACTTAGAGCCAGCACCAAGCTTACTCAGGGCTCCT	13928
Qy	8412	CCGACTTGTAGGAGCTCAGGTTTTCACCCCTTAAACCCAGGGAGCACAGTCCCACTCCA	8471
Db	13927	CCGACTTGTAGGAGCTCAGGTTTTCACCCCTTAAACCCAA-GGAGCACAGTCCCACCTCCA	13869
Qy	8472	CGCCGGGAGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACTTCAGAAATCCA	8531
Db	13868	CGCC-GGGAGCTTAGGACCACTCAGCCCTTAGGAGTATATTTCCGCACTTCAGAAATCCA	13810
Qy	8532	TATCTTGGAATCCAAAGCTCCCTGCCCAATAACTTCAGTCTCTTCCAGAAATTTGGA	8591
Db	13809	TATCTTGGAATCCAAAGCTCCCTGCCCAATAACTTCAGTCTCTTCCAGAAATTTGGA	13750
Qy	8592	AATCCTAGTTTCTCTCTCTCCTC 8612	
Db	13749	TCCCTCTCTCTCTCTCTCTCCTCCTC 13729	
RESULT	6		
LOCUS	AP001187/c		
DEFINITION	AP001187 Homo sapiens chromosome 11 clone RP11-665N17 map 11q13, WORKING DRAFT SEQUENCE, 27 unordered pieces.	HTG	17-JAN-2001
ACCESSION	AP001187		
VERSION	AP001187.3 GI:12246852		
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.		
SOURCE	Homo sapiens DNA, clone:RP11-665N17.		
ORGANISM	Homo sapiens		
REFERENCE	1 (bases 1 to 166269)		
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		

AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Homo sapiens 166,269 genomic DNA of 11q13
 Published Only in DataBase (2000) In press
 2 (bases 1 to 166269)

REFERENCE

AUTHORS

Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 Submitted (18-FEB-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC); Japan
 1-7-22 Suehiro-chou, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
 (E-mail:hattori@gsr.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/
 Tel:81-45-503-9111, Fax:81-45-503-9170)
 On Jan 16, 2001 this sequence version replaced gi:8117633.
 ----- Genome Center

COMMENT

Center: RIKEN Genomic Sciences Center(GSC)
 Center code: RIKEN
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsr.riken.go.jp
 ----- Project Information
 Center Project name: HumDraft11
 Center clone name: RP11-665N17
 ----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 156782 bases at least Q40
 Consensus quality: 160361 bases at least Q30
 Consensus quality: 162264 bases at least Q20
 Insert size: 163669; sum-of-contigs
 Quality coverage: 8.32x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
 27 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved

```

1      28395 contig of 28395 bp in length
      28496      43446 contig of 14951 bp in length
      43447      58083 contig of 14537 bp in length
      58184      66868 contig of 9606 bp in length
      66969      76574 contig of 9606 bp in length
      76675      86418 contig of 9744 bp in length
      86519      94044 contig of 7526 bp in length
      94145      100790 contig of 6846 bp in length
      100891      108539 contig of 7649 bp in length
      108640      116062 contig of 7322 bp in length
      120359      125359 contig of 4197 bp in length
      125460      130186 contig of 5001 bp in length
      130287      134400 contig of 4114 bp in length
      134501      139064 contig of 4564 bp in length
      139165      143906 contig of 4274 bp in length
      144007      148280 contig of 2626 bp in length
      148381      151007 contig of 4727 bp in length
      151107      153748 contig of 4114 bp in length
      153749      153848 contig of 4564 bp in length
      153849      155555 contig of 1707 bp in length
      155556      157199 contig of 1543 bp in length
      157299      159872 contig of 2574 bp in length
      159873      161499 contig of 1527 bp in length
      161500      161599 contig of 100 bp in length
      161600      162703 contig of 1104 bp in length
      162704      162803 contig of 100 bp in length
      162804      163847 contig of 1044 bp in length
      163848      165121 contig of 1174 bp in length
      165122      165221 contig of 1048 bp in length
      165222      166269 contig of 1048 bp in length.
```

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 27 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will

```

* be preserved.
1      28395: contig of 28395 bp in length
      28396      28495: gap of 100 bp
      28496      43446: contig of 14951 bp in length
      43447      43546: gap of 100 bp
      43547      58083: contig of 14537 bp in length
      58084      58183: gap of 100 bp
      58184      66868: contig of 9606 bp in length
      66869      66968: gap of 100 bp
      66969      76574: contig of 9606 bp in length
      76575      76674: gap of 100 bp
      76675      86418: contig of 9744 bp in length
      86419      86518: gap of 100 bp
      86519      94044: contig of 7526 bp in length
      94045      94144: gap of 100 bp
      94145      100790: contig of 6846 bp in length
      100791      100890: gap of 100 bp
      100891      108539: contig of 7649 bp in length
      108540      108639: gap of 100 bp
      108640      115961: contig of 7322 bp in length
      115962      116061: gap of 100 bp
      116062      120258: contig of 4197 bp in length
      120259      120358: gap of 100 bp
      120359      125359: contig of 5001 bp in length
      125360      125459: gap of 100 bp
      125460      130186: contig of 4727 bp in length
      130187      130286: gap of 100 bp
      130287      134400: contig of 4114 bp in length
      134401      134500: gap of 100 bp
      134501      139064: contig of 4564 bp in length
      139065      139164: gap of 100 bp
      139165      143906: contig of 4742 bp in length
      143907      144006: gap of 100 bp
      144007      148280: contig of 4274 bp in length
      148281      148380: gap of 100 bp
      148381      151006: contig of 2626 bp in length
      151007      151106: gap of 100 bp
      151107      153748: contig of 2642 bp in length
      153749      153848: gap of 100 bp
      153849      155555: contig of 1707 bp in length
      155556      155855: gap of 100 bp
      155856      157198: contig of 1543 bp in length
      157199      157298: gap of 100 bp
      157299      159872: contig of 2574 bp in length
      159873      159972: gap of 100 bp
      159973      161499: contig of 1527 bp in length
      161500      161599: gap of 100 bp
      161600      162703: contig of 1104 bp in length
      162704      162803: gap of 100 bp
      162804      163847: contig of 1044 bp in length
      163848      163947: gap of 100 bp
      163948      165121: contig of 1174 bp in length
      165122      165221: gap of 100 bp
      165222      166269: contig of 1048 bp in length.

FEATURES             source
      1..166269
         /organism="Homo sapiens"
         /db_xref="taxon:9606"
         /chromosome="11"
         /map="11q13"
         /clone="RP11-665N17"
      1..28395
         /note="assembly_fragment"
      28496..43446
         /note="assembly_fragment"
      43547..58083
         /note="assembly_fragment"
      58184..66868
         /note="assembly_fragment clone_end:77 vector_side:right"
      66969..76574
         /note="assembly_fragment"
      76675..86418
         /note="assembly_fragment"
```


Db	137738	CCACAGAGTCAATCCCGGTGTGGCTGAGCGGGTATTGTTCCTCCCTCCCCACAGCCTTGTCC	137679
Qy	4535	CTTCATAGTGTAGTAGCCCAAGCCACCCAAAGGACTCCATTTTCTTTGGCCACACCCCTT	4594
Db	137678	CTTCATAGTGTAGTAGCCCAAGCCACCCAAAGGACTCCATTTTCTTTGGCCACACCCCTT	137619
Qy	4595	TCCTCCCATCACCACCCACATAGGAAGGGAAGACAGAGAAGCCCCCTTTCTCGGCTGTCA	4654
Db	137618	TCCTCCCATCACCACCCACATAGGAAGGGAAGACAGAGAAGCCCCCTTTCTCGGCTGTCA	137559
Qy	4655	TTCCCTGAAGCAGGCACAGGGTGGGCCATCATCAGACATAAATCATCTCATCCCCCCTAA	4714
Db	137558	TTCCCTGAAGCAGGCACAGGGTGGGCCATCATCAGACATAAATCATCTCATCCCCCCTAA	137499
Qy	4715	GAGTGGCTGTACCTGAAGGATCATACATGCCTGTGACCGCAAGATGGAGTGGCGTT	4774
Db	137498	GAGTGGCTGTACCTGAAGGATCATACATGCCTGTGACCGCAAGATGGAGTGGCGTT	137439
Qy	4775	CATGGTGTGCCATCAACCCCTTCATTGACCTTGCACACGGACTCCTCTGGAGTTCCTGCA	4834
Db	137438	CATGGTGTGCCATCAACCCCTTCATTGACCTTGCACACGGACTCCTCTGGAGTTCCTGCA	137379
Qy	4835	GCTCAGCAGGTGAGGCTCAGCCAAATGGGGCAGGACTGGGCTAGGCCACAGACTTGACTTG	4894
Db	137378	GCTCAGCAGGTGAGGCTCAGCCAAATGGGGCAGGACTGGGCTAGGCCACAGACTTGACTTG	137319
Qy	4895	CTGTGGGACCTGGGCAAGGGCACTTTCCCTCCTCAGCTTCAGCTTCCTCCCTCGTGA	4954
Db	137318	CTGTGGGACCTGGGCAAGGGCACTTTCCCTCCTCAGCTTCAGCTTCCTCCCTCGTGA	137259
Qy	4955	AATGGGTAGTAATTCCTGGCCTGGCCCTTTCCAGGGCTCTTGGGAGAGTAGAATTGAG	5014
Db	137258	AATGGGTAGTAATTCCTGGCCTGGCCCTTTCCAGGGCTCTTGGGAGAGTAGAATTGAG	137199
Qy	5015	TGTGAATTCGTTTGACTCCATTAAAGGGCTGGTCCCTCAGANAATTTGGCCCTTCCACATGG	5074
Db	137198	TGTGAATTCGTTTGACTCCATTAAAGGGCTGGTCCCTCAGANAATTTGGCCCTTCCACATGG	137139
Qy	5075	TGGGTGGTCCCTGTGGTCTGTACCCCCACCTCTGCCCCATAGGCTTAAGGACCCGTTCTC	5134
Db	137138	TGGGTGGTCCCTGTGGTCTGTACCCCCACCTCTGCCCCATAGGCTTAAGGACCCGTTCTC	137079
Qy	5135	CTCCCTGTTCGGTGGCTCATAACTCTCTCCTTCGGCTCCTTAGAAGCTGCTCTGGCTGCTC	5194
Db	137078	CTCCCTGTTCGGTGGCTCATAACTCTCTCCTTCGGCTCCTTAGAAGCTGCTCTGGCTGCTC	137019
Qy	5195	TATGACCTGGGACATCTGGAAGGTGATGAGGGGAAGTGGCCAGGCTGGGCTGGTGAG	5254
Db	137018	TATGACCTGGGACATCTGGAAGGTGATGAGGGGAAGTGGCCAGGCTGGGCTGGTGAG	136959
Qy	5255	GCCGGGGGCTGGGTGGCAGCTGAATTATGATCTTCTTAGTACCCCATGGCCCTTAG	5314
Db	136958	GCCGGGGGCTGGGTGGCAGCTGAATTATGATCTTCTTAGTACCCCATGGCCCTTAG	136899
Qy	5315	GGAACTTGGCAGATCTTAGAGGAGCTGGAGCCACCCCTTGGCCGGCCAGACCCACTCACCC	5374
Db	136898	GGAACTTGGCAGATCTTAGAGGAGCTGGAGCCACCCCTTGGCCGGCCAGACCCACTCACCC	136839
Qy	5375	TCTACCAACAGTGGGGCATCTAAGGAGGGTGCAGAAGGGAGACCCCTAACAGTGGCTGA	5434
Db	136838	TCTACCAACAGTGGGGCATCTAAGGAGGGTGCAGAAGGGAGACCCCTAACAGTGGCTGA	136779
Qy	5435	GGCAGGGGCCCTCATCTGGGCAGATGAGAGAACTTTGTGTTGGGGGTATCGCCC	5494
Db	136778	GGCAGGGGCCCTCATCTGGGCAGATGAGAGAACTTTGTGTTGGGGGTATCGCCC	136719
Qy	5495	ATTCAGTCTCACTTTGTCTCAACTGTCTGCAGAAATCAGTTCAGTCAAGGCTGTCTCAGGG	5554
Db	136718	ATTCAGTCTCACTTTGTCTCAACTGTCTGCAGAAATCAGTTCAGTCAAGGCTGTCTCAGGG	136659
Qy	5555	GTGTCCAGGGTCCCCAGCCTGGGAGTGGCAGGGGCTGCAATTTGTCCCCCTACGCCCTGCC	5614
Db	136658	GTGTCCAGGGTCCCCAGCCTGGGAGTGGCAGGGGCTGCAATTTGTCCCCCTACGCCCTGCC	136599

Qy	5615	TTTTCTGCCACTGCTTACTGTCTCTTCTCGAGTATACAGAGGTCAATATGTTGGTAGGAC	5674
Db	136598	TTTTCTGCCACTGCTTACTGTCTCTTCTCGAGTATACAGAGGTCAATATGTTGGTAGGAC	136539
Qy	5675	AC---TGAAGAGGGGTGTTCACTTGGTGGGTCTAGGTGGGAGGAGGCCATTTGGGCTG	5731
Db	136538	ACTGATGAAGGGGTGTTCACTTGGTGGGTCTAGGTGGGAGGAGGCCATTTGGGCTG	136479
Qy	5732	GGCTTGAAGTCTTTTGGTGATGTGTAGAAGTGTCTGAGAAGAAGGGCCCTGAGCT	5791
Db	136478	GGCTTGAAGTCTTTTGGTGATGTGTAGAAGTGTCTGAGAAGAAGGGCCCTGAGCT	136419
Qy	5792	CGGAGGCAGGCCCCACCCCTCGAGTCTGCCCCAGGCGCTCAGCCACGCTCTGTAGACC	5851
Db	136418	CGGAGGCAGGCCCCACCCCTCGAGTCTGCCCCAGGCGCTCAGCCACGCTCTGTAGACC	136359
Qy	5852	CAGGAGGAGACCAAGTGAAGGGCTGGCAGCAGTGGAGTGGGAGTGGAGATGGAGAG	5911
Db	136358	CAGGAGGAGACCAAGTGAAGGGCTGGCAGCAGTGGAGTGGGAGTGGAGATGGAGAG	136299
Qy	5912	GACTCCCTGGGATCTTCCCTGTGGCCCTTCTGGGTCTGCCCTGGTCGGGCAATTTGTGCCA	5971
Db	136298	GACTCCCTGGGATCTTCCCTGTGGCCCTTCTGGGTCTGCCCTGGTCGGGCAATTTGTGCCA	136239
Qy	5972	GCAGGGCAGCTGGGGCTGCCCTCGCTGAGGATCCTCTGCCCTCACCTCCATCCAGGGCATTG	6031
Db	136238	GCAGGGCAGCTGGGGCTGCCCTCGCTGAGGATCCTCTGCCCTCACCTCCATCCAGGGCATTG	136179
Qy	6032	CCTCAGCCAGACCTACTATCTGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCT	6091
Db	136178	CCTCAGCCAGACCTACTATCTGGGATGAACACATCTACCCCTACATGTACCTGGCTGGCT	136119
Qy	6092	ACCACTGTGCACCGCAATGTGCGGGAAGCCCTGCAGGCTGGGGCGGACAGGCCACTG	6151
Db	136118	ACCACTGTGCACCGCAATGTGCGGGAAGCCCTGCAGGCTGGGGCGGACAGGCCACTG	136059
Qy	6152	TCATCCAGGAGTAGGATCCCCCTACTTAGGGCTGCAGGCTGTGCTTTCTTCCCCCTCCAT	6211
Db	136058	TCATCCAGGAGTAGGATCCCCCTACTTAGGGCTGCAGGCTGTGCTTTCTTCCCCCTCCAT	135999
Qy	6212	CAGTTTCCAACCCCTGCTCAGGACTGAGGCTTGGCTCCACGCCCCATCCCCCTTCC	6271
Db	135998	CAGTTTCCAACCCCTGCTCAGGACTGAGGCTTGGCTCCACGCCCCATCCCCCTTCC	135939
Qy	6272	ATCCAGTCCCTTAGCACAAGGCCACCATPACCCAGGAGTAGGGACCCCTGATTAGGTG	6331
Db	135938	ATCCAGTCCCTTAGCACAAGGCCACCATPACCCAGGAGTAGGGACCCCTGATTAGGTG	135879
Qy	6332	TCACATCTTCCCTCCCTCCCTCTCTCTCTAAATTTTTTTTTTCTCAGAACAGTCTCAAA	6391
Db	135878	TCACATCTTCCCTCCCTCCCTCTCTCTCTAAATTTTTTTTTTCTCAGAACAGTCTCAAA	135819
Qy	6392	TCTCCATGTTTAAACCACCATCATCCAGCAGTGGGACTTCCACCCCTCGGCCCATGCCCC	6451
Db	135818	TCTCCATGTTTAAACCACCATCATCCAGCAGTGGGACTTCCACCCCTCGGCCCATGCCCC	135759
Qy	6452	CCTCCTCATTTCTTCTTCTCTCTGGGTGACCCAGACAGCATCATTTTGCAGTGGAG	6511
Db	135758	CCTCCTCATTTCTTCTTCTCTCTGGGTGACCCACAGACATCATTTTGCAGTGGAG	135699
Qy	6512	ACCCGACGTACTCCCCCAGGCCCTTGGGGCTCCATCCCCCGCCAGGTCCCTGGGGCTACC	6571
Db	135698	ACCCGACGTACTCCCCCAGGCCCTTGGGGCTCCATCCCCCGCCAGGTCCCTGGGGCTACC	135639
Qy	6572	CCCGATGGTAGACCCCTTCAGACCCCTACAGAGACCCCACTGCTCTCACAGCTACAACCTA	6631
Db	135638	CCCGATGGTAGACCCCTTCAGACCCCTACAGAGACCCCACTGCTCTCACAGCTACAACCTA	135579
Qy	6632	CTGCCGGAAGACGAGGAGATCTACAGAGGTCTTTTGAAGTAGCCAAATGATGCTATCC	6691
Db	135578	CTGCCGGAAGACGAGGAGATCTACAGAGGTCTTTTGAAGTAGCCAAATGATGCTATCC	135519

```
Oy 6692 CAACCTGCTGAAGGAGGAGCCAGCTTGCTGGAGCGGGCGGAGGAGCGCGCGGGGAGCA 6751
|||||
Db 135518 CAACCTGCTGAAGGAGGAGCCAGCTTGCTGGAGCGGGCGGAGGAGCGCGGGGAGCA 135459
|||||
Oy 6752 AAGCCAGGTGAAGGCTGGAGCTCCAGGCTGTGTCCAGGCTCCACCTCCAGGAGGCTCC 6811
|||||
Db 135458 AAGCCAGGTGAAGGCTGGAGCTCCAGGCTGTGTCCAGGCTCCACCTCCAGGAGGCTCC 135399
|||||
Oy 6812 CTTCCACAGGCGCATGGGGCTGATGTACGGGATAGGGATGCGAGGAGGAAGTGCGCC 6871
|||||
Db 135398 CTTCCACA-GGCCATGGGGCTGATGTACGGGATAGGGATGCGAGGAGGAAGTGCGCC 135340
|||||
Oy 6872 CTGAGCAGACAGCTATGTTCCTTTTGTATATACTAGAGTCTCTGGGCCACGTTGGAGCG 6931
|||||
Db 135339 CTGAGCAGACAGCTATGTTCCTTTTGTATATACTAGAGTCTCTGGGCCACGTTGGAGCG 135280
|||||
Oy 6932 GACTGAAGGTATTTAGAGTATTTCTACCTGTGCTTCAGTTTCATGGCCAGACTCCCTC 6991
|||||
Db 135279 GACTGAAGGTATTTAGAGTATTTCTACCTGTGCTTCAGTTTCATGGCCAGACTCCCTC 135220
|||||
Oy 6992 CTTAGCTGAGGGTGGAGGTAGGATGGTACGCTCGCTGCTATGGATGGCTTTTATAAAA 7051
|||||
Db 135219 CTTAGCTGAGGGTGGAGGTAGGATGGTACGCTCGCTGCTATGGATGGCTTTTATAAAA 135160
|||||
Oy 7052 GGAAGAGGTTCTAAGATATGTTCCCAACCTATGCTTACCTTTCTGGAGCCAGGGGCTT 7111
|||||
Db 135159 GGAAGAGGTTCTAAGATATGTTCCCAACCTATGCTTACCTTTCTGGAGCCAGGGGCTT 135100
|||||
Oy 7112 TGCTAGTGGGGGCTGGGCTGTGCCTCTGCTTAAGGGTGTAGTAGACTGATCTG 7171
|||||
Db 135099 TGCTAGTGGGGGCTGGGCTGTGCCTCTGCTTAAGGGTGTAGTAGACTGATCTG 135041
|||||
Oy 7172 TGCCCTCCCTTCCCTCGTCCAGGGCACCCAGAGCCAAAGGTTCCGCCCTCCAGGACCT 7231
|||||
Db 135040 TGCCCTCCCTTCCCTCGTCCAGGGCACCCAGAGCCAAAGGTTCCGCCCTCCAGGACCT 134981
|||||
Oy 7232 GAGTGTCTGGCCCACTCTGCTGATTTACAGAGGCTGTGGGCACTACAGGTGGTGACAGCAGCC 7291
|||||
Db 134980 GAGTGTCTGGCCCACTCTGCTGATTTACAGAGGCTGTGGGCACTACAGGTGGTGACAGCAGCC 134921
|||||
Oy 7292 CCCAGCTGTGCTGACGCTGGGTGGGCCACCTTCTGTGTCAGTCCCTAGGCCGCTTTT 7351
|||||
Db 134920 CCCAGCTGTGCTGACGCTGGGTGGGCCACCTTCTGTGTCAGTCCCTAGGCCGCTTTT 134861
|||||
Oy 7352 GAGGACAGGTGAGGACAGCTGCACAGAGTCTGTGGCACTACAGGTGGTGACAGCAGCC 7411
|||||
Db 134860 GAGGACAGGTGAGGACAGCTGCACAGAGTCTGTGGCACTACAGGTGGTGACAGCAGCC 134801
|||||
Oy 7412 ACGGCTGTGACAGCTTTTCTGGCCAGGGGACGATCTGCCATCCCTTCGGTGGCGGA 7471
|||||
Db 134800 ACGGCTGTGACAGCTTTTCTGGCCAGGGGACGATCTGCCATCCCTTCGGTGGCGGA 134741
|||||
Oy 7472 TGGGACTCAGACCCCTGGGTGGATGGGATGGCCAGAGGAGGTCCTGGAGTTCCAGCC 7531
|||||
Db 134740 TGGGACTCAGACCCCTGGGTGGATGGGATGGCCAGAGGAGGTCCTGGAGTTCCAGCC 134681
|||||
Oy 7532 ACTGCGCGCAACCTTGCTCTCAGCTTGCTCTCCCTCCCTAGGCGCCAGGTCG 7581
|||||
Db 134680 ACTGCGCGCAACCTTGCTCTCAGCTTGCTCTCCCTCCCTAGGCGCCAGGTCG 134631
|||||
```

RESULT 7

```
AP001462
LOCUS AP001462 157454 bp DNA HTG 30-MAY-2000
DEFINITION Homo sapiens chromosome 11 clone RP11-869B15 map 11q13, WORKING
DRAFT SEQUENCE, 33 unordered pieces.
ACCESSION AP001462
VERSION AP001462.2 GI:8117336
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens DNA, clone:RP11-869B15.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
```

REFERENCE

1 (bases 1 to 157454)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Homo sapiens 157,454 genomic DNA of 11q13
Published Only in Database (2000) In press
2 (bases 1 to 157454)
Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
Direct Submission
Submitted (15-MAR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@gsc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)

COMMENT

On May 30, 2000 this sequence version replaced gi:7262597.

----- Genome Center
Center: RIKEN Genomic Sciences Center(GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@gsc.riken.go.jp
----- Project Information
Center project name: Humdraft11
Center clone name: RP11-869B15
----- Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 128493 bases at least Q40
Consensus quality: 142529 bases at least Q30
Consensus quality: 150042 bases at least Q20
Insert size: 154254; sum-of-contigs
Quality coverage: 4.21x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of 33 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved

```
1 20516 contig of 20516 bp in length
20617 30813 contig of 10197 bp in length
30914 42329 contig of 11416 bp in length
42430 53053 contig of 10624 bp in length
53154 63766 contig of 10613 bp in length
63867 73322 contig of 9456 bp in length
73433 81331 contig of 7909 bp in length
81432 88430 contig of 6999 bp in length
88531 93403 contig of 4873 bp in length
93504 98277 contig of 4774 bp in length
98378 101730 contig of 3353 bp in length
101831 106170 contig of 4340 bp in length
106271 110043 contig of 3773 bp in length
110144 114627 contig of 4484 bp in length
114728 117498 contig of 2771 bp in length
117599 120318 contig of 2720 bp in length
120419 122917 contig of 2499 bp in length
123018 125107 contig of 2090 bp in length
125208 127881 contig of 2674 bp in length
127982 130814 contig of 2833 bp in length
130915 132593 contig of 1679 bp in length
132694 134516 contig of 1823 bp in length
134617 137106 contig of 2490 bp in length
137207 140498 contig of 3292 bp in length
140599 143489 contig of 2891 bp in length
143590 145559 contig of 1970 bp in length
145660 147741 contig of 2082 bp in length
147842 149318 contig of 1477 bp in length
149419 150896 contig of 1478 bp in length
150977 152442 contig of 1446 bp in length
152543 154192 contig of 1650 bp in length
154293 156330 contig of 2038 bp in length
```

* 156331 156430: gap of 100 bp	
* 156431 157454: contig of 1024 bp in length.	
FEATURES	Location/Qualifiers
source	1..157454
	/organism="Homo sapiens"
	/db_xref="taxon:9606"
	/chromosome="11"
	/map="11q13"
misc_feature	1..20516
	/note="assembly_fragment"
misc_feature	20617..30813
	/note="assembly_fragment"
misc_feature	30914..42329
	/note="assembly_fragment"
misc_feature	42430..53053
	/note="assembly_fragment"
misc_feature	53154..63766
	/note="assembly_fragment"
misc_feature	63867..73322
	/note="assembly_fragment"
misc_feature	73423..81331
	/note="assembly_fragment"
misc_feature	81432..88430
	/note="assembly_fragment"
misc_feature	88531..93403
	/note="assembly_fragment"
misc_feature	93504..98277
	/note="assembly_fragment"
misc_feature	98378..101730
	/note="assembly_fragment clone_end:SP6 vector_side:right"
misc_feature	101831..106170
	/note="assembly_fragment"
misc_feature	106271..110043
	/note="assembly_fragment"
misc_feature	110144..114627
	/note="assembly_fragment"
misc_feature	114728..117498
	/note="assembly_fragment"
misc_feature	117599..120318
	/note="assembly_fragment"
misc_feature	120419..122917
	/note="assembly_fragment"
misc_feature	123018..125107
Query Match	44.9%; Score 4118.2; DB 82; Length 157454;
Best Local Similarity	94.9%; Pred. No. 0;
Matches 4361; Conservative	0; Mismatches 198; Indels 34; Gaps 15;
Qy 3110	AAGATTATTTTTTTTTTTTTTTTTTTTGTAGACAGGGTCTCACTCTGCCCCAGGCTGGAGT 3169
Db	
48572	ATGTATTTTTTTTATTTTTTTTATTTTTTTTGTAGACGGAGTCTCGTCTGTTGCCAGGCTGGAGT 48631
Qy 3170	GCAGTGGTGTGATCATGGCTCACTGCAGCCTCACTCTCCCTGGGCTCAGAGATCCCTCAA 3229
Db	
48632	ACATGGCGTGATCTCAGCTCACTGCAGAGTCTGCCCTCCGGGTTACGGCATTTCTCTG 48691
Qy 3230	CCTCAGCCTCCTGAGTAGCTCAGAGTACAGGCA---TGCACCGTGGTCTGGTTAAATTT 3286
Db	
48692	CCTCAGCCTCCTGAGTAGCTGGGACTACAGCACTGCCACACACACACACAGCTATTTTTTT 48751
Qy 3287	TTGTATTTTTTTTTTCTAGAGATGGTGTCTCACTATGTGGCCACCACTGGTCTTGAACCTT 3346
Db	
48752	TTTTGTATTTTATAGTAGATGGGGTTTTCACCGTGTAGCCAGGATGGTCTCAATCTCT 48811
Qy 3347	GGGCTCAAGTGATCTGCCCGCCTCAGTCTCCCAA--TGCTGGGATTACAGGTGTGAGCCA 3405
Db	
48812	GACCTCA--TGATCTGCTGCCTCAGCTCAGTCTCCCAAAGTCTGGGATTACAGGTGTGAGCCA 48869
Qy 3406	CCGCAACTGGTGGCCTATGAAA--TTTTTTTTTTTTTTTTCAGACGGGCTCTCACTCTGTG 3464
Db	
48870	CCGCAACTGGTGGCCTATGAAAATTTTTTTTTTTTTTTTTCAGACGGGCTCTCACTCTGTG 48929

Db	51090	CCCTGCGCTTTTCTGCGCAGCTCCTTACTGTCTCCTTGAGTATTAACAGAGTGCTCAATGTG	51149
Qy	5667	GTAGAGCAC---TGAAGAGGGGTGTTCACTTGGTGGGTGTAGTGGGGAGGAGGCCA	5723
Db	51150	GCAGGAGCACTGATGAAGAGGGGTGTTCACTTGGTGGGTGTAGTGGGAGGAGGCCA	51209
Qy	5724	TTGGGCTGGGCTTGAAAGTCTTTGGTGATGTGTAGAAGAGTGTCTGAGAAGAAAGGGC	5783
Db	51210	TTGGGCTGGGCTTGAAAGTCTTTGGTGATGTGTAGAAGAGTGTCTGAGAAGAAAGGGC	51269
Qy	5784	CCTGAGCTCGGAGGGCAGGCCACCCCTGCAGTCTGCCCGAGGCTCAGCCAGGCTCAGCCAGCTCC	5843
Db	51270	CCTGAGCTCGGAGGGCAGGCCACCCCTGCAGTCTGCCCGAGGCTCAGCCAGGCTCAGCCAGCTCC	51329
Qy	5844	TGTAGACCCAGGAGGAGACCAAGGTAGAAGGGCTGGCAGCGAGTGGAGGTGGGAGTGGAG	5903
Db	51330	TGTAGACCCAGGAGGAGACCAAGGTAGAAGGGCTGGCAGCGAGTGGAGGTGGGAGTGGAG	51389
Qy	5904	ATGAGAGGACTCCCTGGGATCTTCTGTGTGGCCCTTCTTGGGTGTGCCCTGGTGGGGCAT	5963
Db	51390	ATGAGAGGACTCCCTGGGATCTTCTGTGTGGCCCTTCTTGGGTGTGCCCTGGTGGGGCAT	51449
Qy	5964	TTGTGCCAGCAGGCGAGCTGGGGCTGCCCTCCCTCAGGATCCTCTGCCTCAGCTCCATCCA	6023
Db	51450	TTGTGCCAGCAGGCGAGCTGGGGCTGCCCTCCCTCAGGATCCTCTGCCTCAGCTCCATCCA	51508
Qy	6024	GGGCATTTGCCCTCAGCAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACTT	6083
Db	51509	GGGCATTTGCCCTCAGCAGACCTACTATCGGGATGAACACATCTACCCCTACATGTACTT	51568
Qy	6084	GGCTGGCTACACATGTCGCAACCCAATGTGCGGGAAGCCCTGCAGGCGCTGGGCGGACAC	6143
Db	51569	GGCTGGCTACACATGTCGCAACCCAATGTGCGGGAAG-CCTGCAGGCGCT-GGCGGACAC	51626
Qy	6144	GGCCACTGTCTATCAGAGTGAGGATCCCCCTACTAGGGCGCTGCAGCGTGTCTTTCTTC	6203
Db	51627	GGCCACTGTCTATCAGAGTGAGGATCCCCCTACTAGGGCGCTGCAGCGTGTCTTTCTTC	51685
Qy	6204	CCCTCCATCAGTTTCCAAACCAACCTCGTCCAGGACTGAGGCGCTGGCTCCCAAGCCCATC	6263
Db	51686	CCCTCCATCAGTTTCCAAACCAACCTCGTCCAGGACTGAGGCGCTGGCTCCCAAGCCCATC	51745
Qy	6264	CCCTTTCCATCCAGTCCCTAGGCAGCAAGGCCACCATTAACCCAGAGGTAGGAGACCCTGA	6323
Db	51746	CCCTTTCCATCCAGTCCCTAGGCAGCAAGGCCACCATTAACCCAGAGGTAGGAGACCCTGA	51805
Qy	6324	TTAAGGTGCACATCTTTCCTCCCTCCCTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	6383
Db	51806	TTAAGGTGCACATCTTTCCTCCCTCCCTCCTCCTCTCTCTCTCTCTCTCTCTCTCTCTCT	51865
Qy	6384	GTCTCAAAATCTCCAAATGTTTAACCAACATATCAGAGCTGGGACTTCCACCCCTCGGCC	6443
Db	51866	GTCTCAAAATCTCCAAATGTTTAACCAACATATCAGAGCTGGGACTTCCACCCCTCGGCC	51925
Qy	6444	CATGCCCCCT	6503
Db	51926	CATGCCCCCT	51985
Qy	6504	CAGTGAGGACCCCACTACTCTCCCGAGCCCTGGGGCTTCCATCCCG-CCAGGTCCCT	6562
Db	51986	CAGTGAGGACCCCACTACTCTCCCGAGCCCTGGGGCTTCCATCCCGCCAGGTCCCT	52045
Qy	6563	GGGGTACCCCGATGTTGAGACCCCTTACAGCCCTACAGAGACCCCACTCTCTCACAG	6622
Db	52046	GGGGTACCCCGATGTTGAGACCCCTTACAGCCCTACAGAGACCCCACTCTCTCACAG	52105
Qy	6623	CTACAACCTACTGCCGGGAAGACGAGGAGATCTACAAGGAGTCTTCTTGAAGTAGCCAATGA	6682
Db	52106	CTACAACCTACTGCCGGGAAGACGAGGAGATCTACAAGGAGTCTTCTTGAAGTAGCCAATGA	52165
Qy	6683	TGTCATCCCAACCTGTGTAAGAGGAGCAGCGCTTGTGGAGGGGGCGGAGAGCGGCC	6742
Db	52166	TGTCATCCCAACCTGTGTAAGAGGAGCAGCGCTTGTGGAGGGGGCGGAGAGCGGCC	52225

QY	6743	GGGGAGCAAGCCAGGTGAAGGCTGGAGCTCCAGCCTGTGTCCAGCCTCCCACTCGGA	6802
Db	52226	GGGGAGCAAGCCAGGTGAAGGCTGGAGCTCCAGCCTGTGTCCAGCCTCCCACTCGGA	52285
QY	6803	CAGGGCTCCCTTCCACAGGGCATGGGGGCTGCATGTACGGGATTTAGGATGGCAGGAGG	6862
Db	52286	CAGGGCTCCCTTCCACA -GGCCATGGGGGCTGCATGTACGGGATTTAGGATGGCAGGAGG	52344
QY	6863	AAGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGGCTATAACTGAGTCTCTGGGCCAC	6922
Db	52345	AAGTGGCCCTGAGCAGACAGCTATGTTCCCTTTTGGCTATAACTGAGTCTCTGGGCCAC	52404
QY	6923	GTTGACGGGACTGAAGTATTTTAGAGGTTTCTACCCCTGTGCCCTTCAGTTTCATGGCCA	6982
Db	52405	GTTGACGGGACTGAAGTATTTTAGAGGTTTCTACCCCTGTGCCCTTCAGTTTCATGGCCA	52464
QY	6983	GACTCCCTCCCTCAGCTCAGGGGTGGAGTAGGATGTACGTCCTGGCTATGGAATTGCC	7042
Db	52465	GACTCCCTCCCTCAGCTCAGGGGTGGAGTAGGATGTACGTCCTGGCTATGGAATTGCC	52524
QY	7043	TTTATAAAGAAAGAGTTCCTAAGATGTTTCCCAACCTATGCTTTACCTTTCTTGGAGCC	7102
Db	52525	TTTATAAAGAAAGAGATTCCTAAGATGTTTCCCAACCTATGCTTTACCTTTCTTGGAGCC	52584
QY	7103	AGGGTCTTTTGGCTAGGTAGTGGGGGCTGGCCTGTGCCCTCTGCTAAGGGGTGAGTAAAG	7162
Db	52585	AGGGCTTTTGGCTTAGTGTGGGGGCTGGCCTGTGCCCTCTGCTAAGGGGTGAGTAAAG	52644
QY	7163	ACTGATCTGTGCCCTCCCTTCCCTCCTGCTCAGGGCACCCAGAGCAAGGTTCCGCCCTC	7222
Db	52645	ACTGATCTGTGCCCTCCCTTCCCTCCTGCTCAGGGCACCCAGAGCAAGGTTCCGCCCTC	52704
QY	7223	CAGGACCTGAGTGTCTCGCCCACTGTCGGATTTCTACGACGGCATCTGCAAAATGGAG	7282
Db	52705	CAGGACCTGAGTGTCTCGCCCACTGTCGGATTTCTACGACGGCATCTGCAAAATGGAG	52764
QY	7283	GAGGCAGTCCCAACGCTGTCTGCACGTGGGCTGGGCACCTTTCTTGTGCAGTCCCTA	7342
Db	52765	GAGGCAGTCCCAACGCTGTCTGCATGTGGGCTGGGCACCTTTCTTGTGCAGTCCCTA	52824
QY	7343	GGCCCTTTTGGGACAGGTGAGGGACAGCTGCACAGAGGCTTGGGCACCTACAGTGGTG	7402
Db	52825	GGCCCTTTTGGGACAGGTGAGGGACAGCTGCACAGAGGCTTGGGCACCTACAGTGGTG	52884
QY	7403	ACAGAGCCACGGGCTGTTCAGACTTTTCTTGSCCCAGGGGACGATCTGCCCATCCCCCT	7462
Db	52885	ACAGAGCCACGGGCTGTTCAGACTTTTCTTGSCCCAGGGGACGATCTGCCCATCCCCCT	52944
QY	7463	CGTGCCGATGGACTGAGACCCCTGGTGGGATGGGATGCCAGACAGGTCCTGGA	7522
Db	52945	GGTGCCGATGGGACTGAGACCCCTGGTGGGATGGGATGGGATGCCAGACAGGTCCTGGA	53004
QY	7523	GTTCCAGGCACCTGGCCGCAACTTGTCTCTCACCTTGTCTCTCCCACTGGCCAGGTGCG	7582
Db	53005	GTTCCAGGCACCTGGCCGCAACTTGTCTCTCACCTTGTCTCTCCCACTGGNNNNNNNN	53064
QY	7583	GCAGAAGTGCCCATAGTAGCCGAGAGGCCGAGGCCGCGGAGGACCGGTGGGG	7642
Db	53065	NN	53124
QY	7643	CGAGGAAGCCGGGAAGCGCGCGGGGCC	7675
Db	53125	NN	53157
RESULT	8		
AP001558			
LOCUS	AP001558	178100 bp	DNA
DEFINITION	AP001558	Homo sapiens chromosome 11 clone RP11-686D5 map liq13, WORKING DRAFT SEQUENCE, 40 unordered pieces.	30-MAY-2000
ACCESSION	AP001558		
VERSION	AP001558.2	GI:8117392	

RESULT	8
AP001558	
LOCUS	
DEFINITION	
ACCESSION	
VERSION	

KEYWORDS

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE

Homo sapiens DNA, clone:RP11-686D5.

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 178100)

REFERENCE

AUTHORS

Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.,
Homo sapiens 178,100 genomic DNA of 11q13
Published Only in Database (2000) In press

TITLE

REFERENCE

2 (bases 1 to 178100)
Hattori, M., Ishii, K., Toyoda, A., Taylor, T.D., Hong-Seog, P.,
Fujiyama, A., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.,
Direct Submission

TITLE

Submitted (29-MAR-2000) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ. 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail: hattori@gsc.riken.go.jp,
URL: http://hgp.gsc.riken.go.jp/, Tel: 81-42-778-9923,
Fax: 81-42-778-9924)

COMMENT

On May 30, 2000 this sequence version replaced gi:7380893.

----- Genome Center

Center: RIKEN Genomic Sciences Center (GSC)

Center code: RIKEN

Web site: http://hgp.gsc.riken.go.jp/

Contact: hattori@gsc.riken.go.jp

----- Project Information

Center project name: Humdraft11

Center clone name: RP11-686D5

----- Summary Statistics

Sequencing vector: PCR products; 100% of reads

Chemistry: Dye-terminator ET-amersham; 100% of reads

Assembly program: Phrap; version 0.990329

Consensus quality: 148042 bases at least Q40

Consensus quality: 162125 bases at least Q30

Consensus quality: 169512 bases at least Q20

Insert size: 174200; sum-of-contigs

Quality coverage: 4.12x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
40 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved

1	20070	contig of	20070 bp	in length
20171	34749	contig of	14579 bp	in length
34850	50049	contig of	15200 bp	in length
50150	60765	contig of	10616 bp	in length
60866	67727	contig of	6862 bp	in length
67828	76362	contig of	8535 bp	in length
76463	82230	contig of	5768 bp	in length
82331	89521	contig of	7191 bp	in length
89622	95129	contig of	5508 bp	in length
95230	101095	contig of	5866 bp	in length
101096	101395	gap of	100 bp	
101196	106255	contig of	5060 bp	in length
106256	106355	gap of	100 bp	
106356	110801	contig of	4446 bp	in length
110802	110901	gap of	100 bp	
110902	114837	contig of	3936 bp	in length
114838	114938	gap of	100 bp	
114939	120430	contig of	5493 bp	in length
120431	120530	gap of	100 bp	
120531	125238	contig of	4708 bp	in length
125239	125338	gap of	100 bp	
125339	128370	contig of	3032 bp	in length
128371	128470	gap of	100 bp	
128471	131184	contig of	2714 bp	in length
131185	131284	gap of	100 bp	
131285	134627	contig of	3343 bp	in length
134628	134727	gap of	100 bp	
134728	137446	contig of	2719 bp	in length
137447	137546	gap of	100 bp	
137547	140281	contig of	2735 bp	in length
140282	140381	gap of	100 bp	
140382	142601	contig of	2220 bp	in length
142602	142701	gap of	100 bp	
142702	145477	contig of	2776 bp	in length
145478	145577	gap of	100 bp	
145578	148565	contig of	2988 bp	in length
148566	148665	gap of	100 bp	
148666	151182	contig of	2517 bp	in length
151183	151282	gap of	100 bp	
151283	153836	contig of	2554 bp	in length
153837	153936	gap of	100 bp	
153937	156236	contig of	2300 bp	in length
156237	156337	contig of	1482 bp	in length

Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 40 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	20070	contig of	20070 bp	in length
20071	20170	gap of	100 bp	
20171	34749	contig of	14579 bp	in length
34750	34849	gap of	100 bp	
34850	50049	contig of	15200 bp	in length
50050	50149	gap of	100 bp	
50150	60765	contig of	10616 bp	in length
60766	60865	gap of	100 bp	
60866	67727	contig of	6862 bp	in length
67728	67827	gap of	100 bp	
67828	76362	contig of	8535 bp	in length
76363	76462	gap of	100 bp	
76463	82230	contig of	5768 bp	in length
82231	82330	gap of	100 bp	
82331	89521	contig of	7191 bp	in length
89522	89621	gap of	100 bp	
89622	95129	contig of	5508 bp	in length
95130	95229	gap of	100 bp	
95230	101095	contig of	5866 bp	in length
101096	101395	gap of	100 bp	
101196	106255	contig of	5060 bp	in length
106256	106355	gap of	100 bp	
106356	110801	contig of	4446 bp	in length
110802	110901	gap of	100 bp	
110902	114837	contig of	3936 bp	in length
114838	114937	gap of	100 bp	
114938	120430	contig of	5493 bp	in length
120431	120530	gap of	100 bp	
120531	125238	contig of	4708 bp	in length
125239	125338	gap of	100 bp	
125339	128370	contig of	3032 bp	in length
128371	128470	gap of	100 bp	
128471	131184	contig of	2714 bp	in length
131185	131284	gap of	100 bp	
131285	134627	contig of	3343 bp	in length
134628	134727	gap of	100 bp	
134728	137446	contig of	2719 bp	in length
137447	137546	gap of	100 bp	
137547	140281	contig of	2735 bp	in length
140282	140381	gap of	100 bp	
140382	142601	contig of	2220 bp	in length
142602	142701	gap of	100 bp	
142702	145477	contig of	2776 bp	in length
145478	145577	gap of	100 bp	
145578	148565	contig of	2988 bp	in length
148566	148665	gap of	100 bp	
148666	151182	contig of	2517 bp	in length
151183	151282	gap of	100 bp	
151283	153836	contig of	2554 bp	in length
153837	153936	gap of	100 bp	
153937	156236	contig of	2300 bp	in length

Qy	4803	GAC	TGCACACCCGACTCGCTGGAGACTTCTG	CAGCTG	CAGCAGGTGAGGGCTGAGCCAA	TG	4862
Db	47918	GAC	TGCACACCCGACTCGCTGGAGACTTCTG	CAGCTG	CAGCAGGTGAGGGCTGAGCCAA	TG	47977
Qy	4863	GGC	CAGGACTGGGCTAGCCAGACTTCAC	TGCTG	TGGACCCCTGGCAGGGGCAC	TTTC	4922
Db	47978	GGC	CAGGACTGGGCTAGCCAGACTTCAC	TGCTG	TGGACCCCTGGCAGGGGCAC	TTTC	48037
Qy	4923	CCT	TCCCTGAGCTTCAGGCTTCCCTTCCT	TGGAAAAATGGGTTAGTAA	TTCTTCTGGCCCTGGCCCT		4982
Db	48038	CCT	TCCCTGAGCTTCAGGCTTCCCTTCCT	TGGAAAAATGGTATAGTAA	TTCTTCTGGCCCTGGCCCT		48097
Qy	4983	TTCC	AGGGCTCTTGGGAGAGTAGAATTGAG	ATGTGAAATTTGCTTTGACTC	CAATTAAGG	5042	
Db	48098	TTCC	AGGGCTCTTGGGAGAGTAGAATTTG	AGATGTGAAATTTGCTTTGACTC	CAATTAAGG	48157	
Qy	5043	GCT	GGTCCCAGAAATTTTGGCCCTTCCAC	ATGGTGGGTGGTCCCTTGGTTG	TCTCACC	5102	
Db	48158	GCT	GGTCCCAGAAATTTTGGCCCTTCCAC	ATGGTGGGTGGTCCCTTGGTTG	TCTCACC	48217	
Qy	5103	ACC	TCTGCCCGATAGGCTAAGGACCCG	TTCTCTCCCTGTTTCGTGGCTCA	TAACTCTCT	5162	
Db	48218	ACC	TCTGCCCGATAGGCTAAGGACCCG	TTCTCTCCCTGTTTCGTGGCTCA	TAACTCTCT	48277	
Qy	5163	CCT	TTCGGCTCCTACAAGCTGCTCTGG	CTGCTATGACCTGGGACATCTG	GAAGGTCAG	5222	
Db	48278	CCT	TTCGGCTCCTACAAGCTGCTCTGG	CTGCTATGACCTGGGACATCTG	GAAGGTCAG	48337	
Qy	5223	TAG	AGGAAGTGGCCAGGCTGCGCTCGG	TGAGGCCGGGGGCTGGGTGGC	AGCCTGAAT	5282	
Db	48338	TAG	AGGAAGTGGCCAGGCTGCGCTCGG	TGAGGCCGGGGGCTGGGTGGC	AGCCTGAAT	48397	
Qy	5283	ATG	ATCCTTTCTAGTTACCCCATGGCC	TTAGGAACTTGGCAGATCTAG	AGGAGCTGGA	5342	
Db	48398	ATG	ATCCTTTCTAGTTACCCCATGGCC	TTAGGAACTTGGCAGATCTAG	AGGAGCTGGA	48457	
Qy	5343	GCC	CACCCCTGGCCGGGCAGACCCAC	TCTACCAAGGTGGGGGCTACT	TAAGGA	5402	
Db	48458	GCC	CACCCCTGGCCGGGCAGACCCAC	TCTACCAAGGTGGGGGCTACT	TAAGGA	48517	
Qy	5403	GGT	TGCAGAGGGAGACCTTAACAGT	GGCTGAGCAGGGGGCCCTCAT	CTGGGCAGATGAG	5462	
Db	48518	GGT	TGCAGAGGGAGACCTTAACAGT	GGCTGAGCAGGGGGCCCTCAT	CTGGGCAGATGAG	48577	
Qy	5463	AAG	AACTTTGTGTGTGGGGGTATCGC	CCATCCAGTCTACCTTTGTG	CAACTGTGT	5522	
Db	48578	AAG	AACTTTGTGTGTGGGGGTATCGC	CCATCCAGTCTACCTTTGTG	CAACTGTGT	48637	
Qy	5523	GC	AAATCAGTTTCTAGTCAAGGCTGT	CTGAGGGGTGTCCAGGGTCC	CCAGCCTGGGAGTG	5582	
Db	48638	GC	AAATCAGTTTCTAGTCAAGGCTGT	CTGAGGGGTGTCCAGGGTCC	CCAGCCTGGGAGTG	48697	
Qy	5583	CG	AGGGCTGCAATTTGTCTCCCTC	TAGCCCTGCTTTCTGCCAC	TGCTTACTTCTCT	5642	
Db	48698	CG	AGGGCTGCAATTTGTCTCCCTC	TAGCCCTGCTTTCTGCCAC	TGCTTACTTCTCT	48757	
Qy	5643	GG	ATATAACAGAGGTCAAATGTGGT	AGGACACCTTGAAGAGGGGTGT	TCCTTGGT	5699	
Db	48758	GG	ATATAACAGAGGTCAAATGTGGT	AGGACACCTTGAAGAGGGGTGT	TCCTTGGT	48817	
Qy	5700	GG	GTAGTGGGAGGGAGGCCATTTGG	CTGGCTGAAAGTCTTTGGT	TCATGTGTAGA	5759	
Db	48818	GG	GTAGTGGGAGGGAGGCCATTTGG	CTGGCTGAAAGTCTTTGGT	TCATGTGTAGA	48877	
Qy	5760	AG	ATGTGTCTGAGAAAGAGAGGGCC	CTGAGCTCGAGGGCAGGCC	CCCCACCCCTGCAGTCT	5819	
Db	48878	AG	ATGTGTCTGAGAAAGAGAGGGCC	CTGAGCTCGAGGGCAGGCC	CCCCACCCCTGCAGTCT	48937	
Qy	5820	CCCC	AGCCCTCAGCCAGCAGTCTCTGT	AGACCCAGGGAGGAGACCG	AGTAAAGGGCTGG	5879	
Db	48938	CCCC	AGCCCTCAGCCAGCAGTCTCTGT	AGACCCAGGGAGGAGACCG	AGTAAAGGGCTGG	48997	
Qy	5880	CAG	CGAGTGGAGGTGGAGTGGAGATG	GAGAGGACTCCCTGGGATCTT	CTCTGTTGGCCCT	5939	

AP000575 73431 bp DNA HTG 30-MAY-2000
LOCUS Homo sapiens chromosome 11 clone XXpl-95F6 map 11q13, WORKING DRAFT
DEFINITION SEQUENCE, 15 unordered pieces.
ACCESSION AP000575
VERSION AP000575.3 GI:8118783
KEYWORDS HTG; HTGS-PHASE1; HTGS-DRAFT.
SOURCE Homo sapiens DNA, clone:XXpl-95F6.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 73431)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Homo sapiens 73,431 genomic DNA of 11q13
JOURNAL Published Only in Database (1999) In press
REFERENCE 2 (bases 1 to 73431)
AUTHORS Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1999) Masahira Hattori, The Institute of Physical
and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
Kitasato Univ., 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
Japan (E-mail:hattori@psc.riken.go.jp,
URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
Fax:81-42-778-9924)
On May 31, 2000 this sequence version replaced gi:6997467.

Center: RIKEN Genomic Sciences Center (GSC)
Center code: RIKEN
Web site: http://hgp.gsc.riken.go.jp/
Contact: hattori@psc.riken.go.jp

Project Information
Center project name: HumDraft11
Center clone name: XXpl-95F6

Summary Statistics
Sequencing vector: PCR products; 100% of reads
Chemistry: Dye-terminator ET-amersham; 100% of reads
Assembly program: Phrap; version 0.990329
Consensus quality: 66753 bases at least Q40
Consensus quality: 69862 bases at least Q30
Consensus quality: 71369 bases at least Q20
Insert size: 72031; sum-of-contigs
Quality coverage: 4.73x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
15 contigs. The true order of the pieces is not known and their
order in this sequence record is arbitrary. Gaps between the
contigs are represented as runs N, but the exact sizes of the gaps
are unknown. This record will be updated with the finished sequence
as soon as it is available and the accession number will be
preserved
1 10174 contig of 10174 bp in length
10275 20667 contig of 10393 bp in length
20768 29453 contig of 8686 bp in length
29554 37145 contig of 7592 bp in length
37246 41949 contig of 4704 bp in length
42050 47344 contig of 5295 bp in length
47445 52652 contig of 5208 bp in length
52753 56084 contig of 3332 bp in length
56185 60854 contig of 4670 bp in length
60955 63499 contig of 2545 bp in length
63600 66501 contig of 2902 bp in length
66602 68545 contig of 1944 bp in length
68646 70748 contig of 2103 bp in length
70849 72288 contig of 1440 bp in length
72389 73431 contig of 1043 bp in length
Sequence updated (12-Oct-1999)
Sequence updated (26-May-2000).
* NOTE: This is a 'working draft' sequence. It currently
consists of 15 contigs. The true order of the pieces
is not known and their order in this sequence record is
arbitrary. Gaps between the contigs are represented as

* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1 10174: contig of 10174 bp in length
10175 10274: gap of 100 bp
10275 20667: contig of 10393 bp in length
20668 20767: gap of 100 bp
20768 29453: contig of 8686 bp in length
29454 29553: gap of 100 bp
29554 37145: contig of 7592 bp in length
37146 37245: gap of 100 bp
37246 41949: contig of 4704 bp in length
41950 42049: gap of 100 bp
42050 47344: contig of 5295 bp in length
47345 47444: gap of 100 bp
47445 52652: contig of 5208 bp in length
52653 52752: gap of 100 bp
52753 56084: contig of 3332 bp in length
56085 56184: gap of 100 bp
56185 60854: contig of 4670 bp in length
60855 60954: gap of 100 bp
60955 63499: contig of 2545 bp in length
63500 63599: gap of 100 bp
63600 66501: contig of 2902 bp in length
66502 66601: gap of 100 bp
66602 68545: contig of 1944 bp in length
68546 68645: gap of 100 bp
68646 70748: contig of 2103 bp in length
70749 70848: gap of 100 bp
70849 72288: contig of 1440 bp in length
72289 72388: gap of 100 bp
72389 73431: contig of 1043 bp in length.

FEATURES
Location/Qualifiers
1..73431
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="XXpl-95F6"
1..10174
/note="assembly_fragment"
10275..20667
/note="assembly_fragment"
20768..29453
/note="assembly_fragment"
29554..37145
/note="assembly_fragment"
37246..41949
/note="assembly_fragment"
42050..47344
/note="assembly_fragment"
47445..52652
/note="assembly_fragment"
52753..56084
/note="assembly_fragment"
56185..60854
/note="assembly_fragment"
60955..63499
/note="assembly_fragment"
63600..66501
/note="assembly_fragment"
66602..68545
/note="assembly_fragment"
68646..70748
/note="assembly_fragment"
70849..72288
/note="assembly_fragment"
72389..73431
/note="assembly_fragment"
BASE COUNT 17276 a 18750 c 19264 g 16741 t 1400 others
ORIGIN

Query Match	22.6%	Score 2071.4	DB 82	Length 73431
Best Local Similarity	94.3%	Pred. No. 0		
Matches 2098	Conservative	0	Mismatches 126	Indels 1
Gaps				1
QY	1044	AATATATATATATATATATATGTGTCGTGTGTGTGTGTGTATATATATATATA	1103	
DB	68647	ATATATATATATATATGTGTCGTGTGTGTGTGTGTGTGTGTGTATATATATATATA	68706	
QY	1104	TATTTATGTATATGCACATACACAAAAAATAGCGGGAGTGTTGGCCACGCCCTGTGAT	1163	
DB	68707	TATTTATGTATATGCACATACACAAAAAATAGCGGGAGTGTTGGCCACGCCCTGTGAT	68766	
QY	1164	CACAGCTACTCGGAGGCTGAGGCACGAGATCCTTCAGCCCTGTAAGTCGAGCGCTGCA	1223	
DB	68767	CACAGCTACTCGGAGGCTGAGGCACGAGATCCTTCAGCCCTGTAAGTCGAGCGCTGCA	68826	
QY	1224	GTGAGCCAGATCGAGGCACCTGCATTCAGCCTGGGGAAAGAGAAGAACCGTGTCTCAA	1283	
DB	68827	GTGAGCCAGATCGAGGCACCTGCATTCAGCCTGGGGAAAGAGAAGAACCGTGTCTCAA	68886	
QY	1284	AACAACAAACAAGCTACTCTTAGCACGTGTGTAGAGTATCTCGCGGGCGGAAGTGGGA	1343	
DB	68887	AACAACAAACAAGCTACTCTTAGCACGTGTGTAGAGTATCTCGCGGGCGGAAGTGGGA	68946	
QY	1344	AACCAGTGCTGCACACAGAGTAGGCATCTTTATATGTTTAAACAGACACTGATACCCAGCTA	1403	
DB	68947	AACCAGTGCTGCACACAGAGTAGGCATCTTTATATGTTTAAACAGACACTGATACCCAGCTA	69006	
QY	1404	AAGCGGTGAACACATTTACTCTCTGCGACGTGTTTTAAAAGTATCTGTTTTCTCATATPG	1463	
DB	69007	AAGCGGTGAACACATTTACTCTCTGCGACGTGTTTTAAAAGTATCTGTTTTCTCATATPG	69066	
QY	1464	TTTTTATTTTAAATTTTCTGGATCAAGCAACTGATCTTTTTTCCCTCATAACTGCCGACC	1523	
DB	69067	TTTTTATTTTAAATTTTCTGGATCAAGCAACTGATCTTTTTTCCCTCATAACTGCCGACC	69126	
QY	1524	GACCCGTGACAGCAAAAACCGGCAGAAGCTCGGCAGACCTCCCACCCCGAGTCTGCAGGTAG	1583	
DB	69127	GACCCGTGACAGCAAAAACCGGCAGAAGCTCGGCAGACCTCCCACCCCGAGTCTGCAGGTAG	69186	
QY	1584	TGCCCCCGGACTACATTTTCCAGAAGGCACCTGCGGGACAGCTTCTGCTCGTGGCGCTG	1643	
DB	69187	TGCCCCCGGACTACATTTTCCAGAAGGCACCTGCGGGACAGCTTCTGCTCGTGGCGCTG	69246	
QY	1644	AAGGGAAGGGCCAATCCCTGAGTATCTCGGGAAGGAGGTGTCGGAGCGCGGGACCTAGA	1703	
DB	69247	AAGGGAAGGGCCAATCCCTGAGTATCTCGGGAAGGAGGTGTCGGAGCGCGGGACCTAGA	69306	
QY	1704	GATCCAGAACCCACAGGCGAGCGGCCCGGCCGCTATTTCCAGGCTCTGGGGGCA	1763	
DB	69307	GATCCAGAACCCACAGGCGAGCGGCCCGGCCGCTATTTCCAGGCTCTGGGGGCA	69366	
QY	1764	GGGTGGGCCCCAGACTCCACTTCCCGGGGGTAGTGCAGACCTAGGGGCGGACTTCAATG	1823	
DB	69367	GGGTGGGCCCCAGACTCCACTTCCCGGGGGTAGTGCAGACCTAGGGGCGGACTTCAATG	69426	
QY	1824	TCCAGCAGGCTCCGGCGCGGTGCGCGCGCGGTGCCTAGTGTGGATGTAAGCGCGGAGG	1883	
DB	69427	TCCAGCAGGCTCCGGCGCGGTGCGCGCGCGGTGCCTAGTGTGGATGTAAGCGCGGAGG	69486	
QY	1884	TGGGGAGGGGACCGAGGCCAGGACTCTCCTTGGGGTTTGGGGCTTTCACCTGGGTGGG	1943	
DB	69487	TGGGGAGGGGACCGAGGCCAGGACTCTCCTTGGGGTTTGGGGCTTTCACCTGGGTGGG	69546	
QY	1944	CTTTCTGGACAGACTTTTACAGCCCGGGGGCACAGTCTGTAGAGAGGGGGGGCGGCC	2003	
DB	69547	CTTTCTGGACAGACTTTTACAGCCCGGGGGCACAGTCTGTAGAGAGGGGGGGCGGCC	69605	
QY	2004	ATTGGGGTCTCTCATTTGGGGTGTCTTGGGGCGCACCCCATCATCGGGTACCGGGCGTCCCGGAA	2063	
DB	69606	ATTGGGGTCTCTCATTTGGGGTGTCTTGGGGCGCACCCCATCATCGGGTACCGGGCGTCCCGGAA	69665	

QY	2064	TTGTGGGGACAAAAGGCTCTGCAGTCTCGGCTGAGGGGTCTCACCAACAAAGAGGGG	2123
Db	69666	TTGTGGGGACAAAAGGCTCTGCAGTCTCGGCTGAGGGGTCTCACCAACAAAGAGGGG	69725
QY	2124	AAGCCGGTGACAGAGGCTGAAGAGGTTGGGAACGACGGAGCTGTCCGTGTGTGTCGGGG	2183
Db	69726	AAGCCGGTGACAGAGGCTGAAGAGGTTGGGAACGACGGAGCTGTCCGTGTGTGTCGGGG	69785
QY	2184	CGGTTGGAACCTTAGCGGACCTTGGAGGAGGCTCCCGCGCCGAAACCTGCCGACCCCTCC	2243
Db	69786	CGGTTGGAACCTTAGCGGACCTTGGAGGAGGCTCCCGCGCCGAAACCTGCCGACCCCTCC	69845
QY	2244	CTCCCGCCGGCTTGCCTTCAGAGCGCCGCCACCGCCGCCCATGGGCTGAAGGGCG	2303
Db	69846	CTCCCGCCGGCTTGCCTTCAGAGCGCCGCCACCGCCGCCCATGGGCTGAAGGGCG	69905
QY	2304	CCGAAGACGCTGTTCGCCGTGCGCTCCATCGAGAGCTGGTGGCGCTGTGTGCTGGCG	2363
Db	69906	CCGAAGACGCTGTTCGCCGTGCGCTCCATCGAGAGCTGGTGGCGCTGTGTGCTGGCG	69965
QY	2364	AGCTGGGCGGAGAGAGCGGACCTGGTGTCTTTCCTTGGTGTGGGCTTCGTGGAGC	2423
Db	69966	AGCTGGGCGGAGAGAGCGGACCTGGTGTCTTTCCTTGGTGTGGGCTTCGTGGAGC	70025
QY	2424	ATTTTCTGGCTGTCAAACCGCGTCATCCCTACCAAGCTTCCCGAGCTCACCTTCCAGCCCA	2483
Db	70026	ATTTTCTGGCTGTCAAACCGCGTCATCCCTACCAAGCTTCCCGAGCTCACCTTCCAGCCCA	70085
QY	2484	GCCCGCCCGGACCCCGCTGGCGGCTCACCTACTTTCCTGGTGGCGACCTGTCTATCA	2543
Db	70086	GCCCGCCCGGACCCCGCTGGCGGCTCACCTACTTTCCTGGTGGCGACCTGTCTATCA	70145
QY	2544	TCGCCGCCCTCTATGCCCGCTTTCACCGCCAGATCCGAGGGCGCTGCACCTGTCCCTCT	2603
Db	70146	TCGCCGCCCTCTATGCCCGCTTTCACCGCCAGATCCGAGGGCGCTGCACCTGTCCCTCT	70205
QY	2604	ATCTCTGAGAAGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAAGTCTCCCGATGTCATAT	2663
Db	70206	ATCTCTGAGAAGGGGTGTCTCCAGCCGTGAGCTGGTGAAGAAGTCTCCCGATGTCATAT	70265
QY	2664	GGAACAGCTCAGCCGCTCTACTTCAAGGATCGGGGCCACATCCAGTCCCTCTTCAGCT	2723
Db	70266	GGAACAGCTCAGCCGCTCTACTTCAAGGATCGGGGCCACATCCAGTCCCTCTTCAGCT	70325
QY	2724	TCATCACAGTTTGGAGCCGAGTAGTGGAATCTTTATCCATGACCCACTTCTTCAAAACC	2783
Db	70326	TCATCACAGTTTGGAGCCGAGTAGTGGAATCTTTATCCATGACCCACTTCTTCAAAACC	70385
QY	2784	CTCCATGTTTACAGAACCTTTTAAGAACTGTAAAGCTTGTGAGTTCGGCAGCTGTTA	2843
Db	70386	CTCCATGTTTACAGAACCTTTTAAGAACTGTAAAGCTTGTGAGTTCGGCAGCTGTTA	70445
QY	2844	TTTTTCTCTTTTCAGTTTGGGAACTGAAGCCAGAGAGGGGAAATGATATGCCAAAGTCA	2903
Db	70446	TTTTTCTCTTTTCAGTTTGGGAACTGAAGCCAGAGAGGGGAAATGATATGCCAAAGTCA	70505
QY	2904	CACACGGATGCGAGGGCTGGAAGTGAAGCTTGATCACTTGGCTCCAAATCATCAACCTC	2963
Db	70506	CACACGGATGCGAGGGCTGGAAGTGAAGCTTGATCACTTGGCTCCAAATCATCAACCTC	70565
QY	2964	ACCTCTGCCCTCTAGCAGCCGCCACCTTGGCCACTGAACAGCTACAGGATTTCTAAGCAT	3023
Db	70566	ACCTCTGCCCTCTAGCAGCCGCCACCTTGGCCACTGAACAGCTACAGGATTTCTAAGCAT	70625
QY	3024	GAGACACAGAGGGCGGACAGATTTAGGGGCAAGAAGATGAAATTTGGGCTGCATTTGA	3083
Db	70626	GAGACACAGAGGGCGGACAGATTTAGGGGCAAGAAGATGAAATTTGGGCTGCATTTGA	70685
QY	3084	GCGAGTTAAACAAAATAATGGCTATCAAGATTTTTTTTTTTTTTTTTTTTTTTTTCAGACGG	3143
Db	70686	GCGAGTTAAACAAAATAATGGCTATCAAGATTTTTTTTTTTTTTTTTTTTTTTTTTTTGA	70745
QY	3144	TCCTACTCTGTCCCGCAGGCTGGAGTGGTGTGATCATGGCTCACTGCAGGCTCAG	3203

108640	115961	contig of	7322	bp in length
116062	120258	contig of	4197	bp in length
120359	125359	contig of	5001	bp in length
125460	130186	contig of	4727	bp in length
130287	134400	contig of	4114	bp in length
134501	139064	contig of	4564	bp in length
139165	143906	contig of	4742	bp in length
144007	148280	contig of	4274	bp in length
148381	151006	contig of	2626	bp in length
151107	153748	contig of	2642	bp in length
153849	155555	contig of	1707	bp in length
155656	157198	contig of	1543	bp in length
157299	159872	contig of	2574	bp in length
159973	161499	contig of	1527	bp in length
161600	162703	contig of	1104	bp in length
162804	163847	contig of	1044	bp in length
163948	165121	contig of	1174	bp in length
165222	166269	contig of	1048	bp in length.

* NOTE: This is a 'working draft' sequence. It currently
 * consists of 27 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved.

1	28395;	contig of	28395	bp in length
28396	28495;	gap of	100	bp
28496	43446;	contig of	14951	bp in length
43447	43546;	gap of	100	bp
43547	58083;	contig of	14537	bp in length
58084	58183;	gap of	100	bp
58184	66868;	contig of	8685	bp in length
66869	66968;	gap of	100	bp
66969	78574;	contig of	3606	bp in length
78575	76674;	gap of	100	bp
76675	86418;	contig of	9744	bp in length
86419	86518;	gap of	100	bp
86519	94044;	contig of	7526	bp in length
94045	94144;	gap of	100	bp
94145	100790;	contig of	6646	bp in length
100791	100890;	gap of	100	bp
100891	108539;	contig of	7649	bp in length
108540	108639;	gap of	100	bp
108640	115961;	contig of	7322	bp in length
115962	116061;	gap of	100	bp
116062	120258;	contig of	4197	bp in length
120259	120358;	gap of	100	bp
120359	125359;	contig of	5001	bp in length
125360	125459;	gap of	100	bp
125460	130186;	contig of	4727	bp in length
130187	130286;	gap of	100	bp
130287	134400;	contig of	4114	bp in length
134401	134500;	gap of	100	bp
134501	139064;	contig of	4564	bp in length
139065	139164;	gap of	100	bp
139165	143906;	contig of	4742	bp in length
143907	144006;	gap of	100	bp
144007	148280;	contig of	4274	bp in length
148281	148380;	gap of	100	bp
148381	151006;	contig of	2626	bp in length
151007	151106;	gap of	100	bp
151107	153748;	contig of	2642	bp in length
153749	153848;	gap of	100	bp
153849	155555;	contig of	1707	bp in length
155556	155655;	gap of	100	bp
155656	157198;	contig of	1543	bp in length
157199	157298;	gap of	100	bp
157299	159872;	contig of	2574	bp in length
159873	159972;	gap of	100	bp
159973	161499;	contig of	1527	bp in length
161500	161599;	gap of	100	bp
161600	162703;	contig of	1104	bp in length

* 162704 162803: gap of 100 bp
* 162804 163847: contig of 1044 bp in length
* 163848 163947: gap of 100 bp
* 163948 165121: contig of 1174 bp in length
* 165122 165221: gap of 100 bp
* 165222 166269: contig of 1048 bp in length.

FEATURES

source

Location/Qualifiers
1. .166269
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="11"
/map="11q13"
/clone="RP11-665N17"
1. .28395
/note="assembly_fragment"
28496. .43446
/note="assembly_fragment"
43547. .58083
/note="assembly_fragment"
58184. .66868
/note="assembly_fragment clone_end:T7 vector_side:right"
66969. .76574
/note="assembly_fragment"
76675. .86418
/note="assembly_fragment"
86519. .94044
/note="assembly_fragment"
94145. .100790
/note="assembly_fragment"
100891. .108539
/note="assembly_fragment"
108640. .115961
/note="assembly_fragment"
116062. .120258
/note="assembly_fragment clone_end:SP6 vector_side:left"
120359. .125359
/note="assembly_fragment"
125460. .130186
/note="assembly_fragment"
130287. .134400
/note="assembly_fragment"
134501. .139064
/note="assembly_fragment"
139165. .143906
/note="assembly_fragment"
144007. .148280
/note="assembly_fragment"
148381. .151006
/note="assembly_fragment"
151107. .153748
/note="assembly_fragment"
153849. .155555
/note="assembly_fragment"
155656. .157198
/note="assembly_fragment"
157299. .159872
/note="assembly_fragment"
159973. .161499
/note="assembly_fragment"
161600. .162703
/note="assembly_fragment"
162804. .163847
/note="assembly_fragment"
163948. .165121
/note="assembly_fragment"
165222. .166269
/note="assembly_fragment"
37933 a 44411 c 44369 g 36956 t 2600 others

Query Match 21.6%; Score 1982.2; DB 82; Length 166269;
Best Local Similarity 93.9%; Pred. No. 0;
Matches 2054; Conservative 0; Mismatches 128; Indels 5; Gaps 5;

QY 1 CTGGTCTTGAACCTCCGGCCCTCAAGCAATCCTCCTGCTTCCAGCTTCCCAAGTGTCTTAA 60
Db 146200 CTGGTCTTGAACCTCCGGCCCTCAAGCAATCCTCCTGCTTCCAGCTTCCCAAGTGTCTTAA 146259
QY 61 TTACAGCCATGACCCCTGGCCATGAACCTTGACACATATCAGATATACCTGCTCAGGTATTTTG 120
Db 146260 TTACAGCCATGACCCCTGGCCATGAACCTTGACACATATCAGATATACCTGCTCAGGTATTTTG 146319
QY 121 TGGAAATGTCCTCAACTCTCTTTTGGCCAGATGTTTCTCATGATTAGAGGAGAGTATAA 180
Db 146320 TGGAAATGTCCTCAACTCTCTTTTGGCCAGATGTTTCTCATGATTAGAGGAGAGTATAA 146379
QY 181 ATTTGAGGAAAATCCAGAGAGGTGAAGAGGTGAAGTGAAGGAGGAGGAGGAGGAGGAGGAGG 240
Db 146380 ATTTGAGGAAAATCCAGAGAGGTGAAGAGGTGAAGTGAAGGAGGAGGAGGAGGAGGAGG 146439
QY 241 ATTTACTGCTATATACCGAGTGTCTGGAACCTTGCCCATGCTGTAAGTACCAAAAATCTGTT 300
Db 146440 ATTTACTGCTATATACCGAGTGTCTGGAACCTTGCCCATGCTGTAAGTACCAAAAATCTGTT 146499
QY 301 TTTTGTGAATGAATAAGCAAAATAAATGAGTGACCGTGGAAATTTAGTATTATTTTCAAAGT 360
Db 146500 TTTTGTGAATGAATAAGCAAAATAAATGAGTGACCGTGGAAATTTAGTATTATTTTCAAAGT 146559
QY 361 TTCAAAGCGTTTGTGATACAGGCCAGGCAGTGGCTCACACCTGTAAATCCAGCACCTTT 420
Db 146560 TTCAAAGCGTTTGTGATACAGGCCAGGCAGTGGCTCACACCTGTAAATCCAGCACCTTT 146619
QY 421 TGGAGCCGAGGTAGGAGGATCACTTCAGGTCAGGAGTTCGAGACCAAGCTTGACCAACAT 480
Db 146620 TGGAGCCGAGGTAGGAGGATCACTTCAGGTCAGGAGTTCGAGACCAAGCTTGACCAACAT 146679
QY 481 GGTGACACCCCTGTCTCTACTTAAGTAAATAACAAAATTAAGCCAAAGTGTGGTGGCAGGCA 540
Db 146680 GGTGACACCCCTGTCTCTACTTAAGTAAATAACAAAATTAAGCCAAAGTGTGGTGGCAGGCA 146739
QY 541 CCTGTAATCCGGCTACTTGGGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 600
Db 146740 CCTGTAATCCGGCTACTTGGGAAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 146799
QY 601 AGGTTGCAGTGAGCCGAGATCACCCACTGCACCTCCAGCCTGAGTGACAGAGCAGAGACTC 660
Db 146800 AGGTTGCAGTGAGCCGAGATCACCCACTGCACCTCCAGCCTGAGTGACAGAGCAGAGACTC 146859
QY 661 TGCTCTCAAAAACAAATAAACAATAACTACTCTTTGGCCGGGTAAAGTGGTTCAGCCCTGT 720
Db 146860 TGCTCTCAAAAACAAATAAACAATAACTACTCTTTGGCCGGGTAAAGTGGTTCAGCCCTGT 146919
QY 721 AATTTAGCACTTTGGGAGGCTGAGGCGGCGGAGATCACTTGAGGTTAGGGTTTCGAGACC 780
Db 146920 AATTTAGCACTTTGGGAGGCTGAGGCGGCGGAGATCACTTGAGGTTAGGGTTTCGAGACC 146979
QY 781 AGCTGGCCCAACATGGTGAACCCCATCTCTACTTAAATACAAAAGTTTCTCGGGTGT 840
Db 146980 AGCTGGCCCAACATGGTGAACCCCATCTCTACTTAAATACAAAAGTTTCTCGGGTGT 147039
QY 841 GGTGGCGGAGCCCTATAATCCAGCTACTTTGGGACTTTTTTTTAAAGACGGAATCTCACTC 900
Db 147040 GGTGGCGGAGCCCTATAATCCAGCTACTTTGGGACTTTTTTTTAAAGACGGAATCTCACTC 147099
QY 901 TGTGTCAGGCTGGAGTGAGTGGCAAGATCTTGGCTCACTGAAGCCTCGGCCCTCCAG 960
Db 147100 TGTGTCAGGCTGGAGTGAGTGGCAAGATCTTGGCTCACTGAAGCCTCGGCCCTCCAG 147159
QY 961 GTTCAAGGGATTCCCGCGCTCAGCCTCCCAAGTGAAGTGGGAATCCCTCTCTCTGCAA 1020
Db 147160 GTTCAA -GGGATTCGCGCGCTCAGCCTCCCAAGTGAAGTGGGAATCCCTCTCTCTGCAA 147218
QY 1021 AAAAAAAAAAAAAAAAAAAAAAAT 1080
Db 147219 AAAAAAAAAAAAAAAAAAAAAAAT 147278
QY 1081 TGTGTGTGTGTAT 1140

Db	147279	TGTTGGTGTATTATGATATATATATTTATGTATATGCACATACACAAAAATTAGCGGG	147338
Qy	1141	GAGTGGTGGCGACCGCTGTGATCACAGCTACTCGGAGGCTCAGGCACGAGAAATCGCTT	1200
Db	147339	GAGTGGTGGCGACCGCTGTGATCACAGCTACTCGGAGGCTCAGGCACGAGAAATCGCTT	147398
Qy	1201	GAGCCGTGAAGTCGAGGCTGCAGTGAAGCCAGATCGAGCACTGCATCTCCAGCCTGGCG	1260
Db	147399	GAGCCGTGAAGTCGAGGCTGCAGTGAAGCCAGATCGAGCACTGCATCTCCAGCCTGGCG	147458
Qy	1261	GAAGAGAAAGACCGTGTCTCAAAACAAACAAACAAAGCTACTCTAGCAGTGTTAGA	1320
Db	147459	GAAGAGAAAGACCGTGTCTCAAAACAAACAAACAAAGCTACTCTAGCAGTGTTAGA	147518
Qy	1321	GTATCTCGCGGGCGGAAGTGGGAAACGAGTGTCTCAACAGAGTAGGCACTCTTATATGT	1380
Db	147519	GTATCTCGCGGGCGGAAGTGGGAAACGAGTGTCTCAACAGAGTAGGCACTCTTATATGT	147578
Qy	1381	TAAACAGACTGATACCCAGCTAAAGCGGCTGAACACATTTACTCTCGCAGTGTTTAA	1440
Db	147579	TAAACAGACTGATACCCAGCTAAAGCGGCTGAACACATTTACTCTCGCAGTGTTTAA	147638
Qy	1441	AAGTATCTGTTTTCTCATATGTGTTTTTAAATTTTTCTCGATCAGCAACCTGATC	1500
Db	147639	AAGTATCTGTTTTCTCATATGTGTTTTTAAATTTTTCTCGATCAGCAACCTGATC	147698
Qy	1501	TTTTTCTCATAACTTCCCGACCGACCGTGCAGAGCAAAACCGGCAGAGCTCGCGCAAC	1560
Db	147699	TTTTTCTCATAACTTCCCGACCGACCGTGCAGAGCAAAACCGGCAGAGCTCGCGCAAC	147757
Qy	1561	TCCACACCCGAGTCTCAGAGTAGTCCCGCGGACTACATTTCCAGAGCACTGCGGCG	1620
Db	147758	TCCACACCCGAGAGTCTCAGAGTAGTCCCGCGGACTACATTTCCAGAGCACTGCGGCG	147817
Qy	1621	ACGCTTCTGCTGCTCGGCTGAAAGGAAGGCGCAATCCTCAGTATCTCGGGAAGAG	1680
Db	147818	ACGCTTCTGCTGCTCGGCTGAAAGGAAGGCGCAATCCTCAGTATCTCGGGAAGAG	147877
Qy	1681	GTGTCGAGAGCCGCGACCTAGAGATCCAGAACCCACAGCGAGCGGCCCGC -CCGCG	1739
Db	147878	GTGTCGAGAGCCGCGAGCTAGAGATCCAGAACCCACAGCGAGCGGCCCGC -CCGCG	147937
Qy	1740	ACTATTTCCAGGCTCTCGGGGCGAGGGTGGGCCAGACTCCACTTCCGGCGGGTAGTG	1799
Db	147938	ACTATTTCCAGGCTCTCGGGGCGAGGGTGGGCCAGACTCCACTTCCGGCGGGTAGTG	147997
Qy	1800	CGACCTTAGGGCGGGGACTTCATGTCACAGCGCTCGGGCGGCTGCGCGCGGTGCC	1859
Db	147998	CGACCTTAGGGCGGGTACTTCATGTCACAGCGCTCGGGCGGCTGCGCGCGGTGCC	148057
Qy	1860	TAGTGTGGGATGTTAAGCGCGAGGTGGGCGAGGGGACCGAGCCAGGACTCTCCTTGGG	1919
Db	148058	TAGTGTGGGATGTTAAGCGCGAGGTGGGCGAGGGGACCGAGCCAGGACTCTCCTTGGG	148117
Qy	1920	GTTTGGGGCTTGACCTGGGTGCGCTTCTGTGACAGACTTTACAGCCCGCGGGGACAG	1979
Db	148118	GTTTGGGGCTTGACCTGGGTGCGCTTCTGTGACAGACTTTACAGCCCGCGGGGACAG	148176
Qy	1980	TCGTAGAGAGGGGCGGGCGGCATTGGGGCTCTCTATTGGGGTGTCTGGGCGCACCC	2039
Db	148177	TCGTAGAGAGGGGCGGGCTCGGCCATTGGGGCTCTCTATTGGGGTGTCTGGGCGCACCC	148236
Qy	2040	CATCGGTACCGGGCTCCCGGAATGTGGGGGACAAAAAGGCTCTGCACTCTCGGCTGA	2099
Db	148237	CATCGGTACCGGGCTCCCG -GAATGTGGGGGACAAAAAGGCTCTCGGCTCTCGGCTGA	148295
Qy	2100	GGGTCTCACCGACAAAGAGGGGACCGCTGAGCAGAGGCTGAAGAGGTGCGGAAC	2159
Db	148296	NN	148355
Qy	2160	AGGGAGAGCTGTGCTGTGCGGGCGGG	2186

Db 148356 NNNNNNNNNNNNNNNNNNNNNNNCNG 148382

RESULT 11

HSU93236 2772 bp mRNA PRI 19-APR-1997

LOCUS Human menin (MEN1) mRNA, complete cds.

DEFINITION U93236

ACCESSION U93236

VERSION U93236.1 GI:1945386

KEYWORDS

SOURCE human.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo. 1 (bases 1 to 2772)

AUTHORS Chandrasekharappa,S.C., Guru,S.C., Manickam,P., Olufemi,S.-E., Collins,F.S., Emmert-Luck,M.R., Debelenko,L.V., Zhuang,Z., Lubensky,I.A., Liotta,L.A., Crabtree,J.S., Wang,Y., Roe,B.A., Weismann,J., Boguski,M.S., Agarwal,S.K., Kester,M., Kim,Y.S., Heppner,C., Dong,Q., Spiegel,A.M., Burns,A.L. and Marx,S.J.

TITLE Positional cloning of the gene for multiple endocrine neoplasia-type 1

JOURNAL Science 276 (5311), 404-407 (1997)

MEDLINE 97258940

REFERENCE 2 (bases 1 to 2772)

AUTHORS Collins,F.S.

TITLE Direct Submission

JOURNAL Submitted (13-MAR-1997) National Human Genome Research Institute, Bldg 38A, Room 605, National Institutes of Health, Bethesda, MD 20892, USA

FEATURES

source Location/Qualifiers

1..2772

/organism="Homo sapiens"

/db_xref="taxon:9606"

/chromosome="11"

/map="11q13"

1..87

/gene="MEN1"

/number=1

1..2772

/gene="MEN1"

88..555

/gene="MEN1"

/number=2

111..1943

/gene="MEN1"

/codon_start=1

/product="menin"

/protein_id="AAC5128.1"

/db_xref="GI:1945387"

/translation="MGLKAAQKTLFLRSIDDDVRLFAAEELGKEPDLVLLSLVLGVF EHLAVNRVITPVNPKTVTFQSPAPPPGGTLTFPVADLSI IAAIVARFTAOIRGAVFA LSLYPREGGVSSRELKVKVSDVILNLSRSYFKDRAHIQSLFSFICTKLDSGGAFA VVACAGALGRDVLHALSEDHANVYRPGNCEQTAEVTHGKGNEDRRGQTVNAGVAFR SWILYKSGDRCKMEVAFMVCATINPSIDLHFDLSLELOQKRLWLLYDUGHLERY PMALGNLWADTATYIQDNYCREDEEYKEFFEVANDVIYIYRDEHIYPIYMLAGYHCRNRY REALQWADTATYIQDNYCREDEEYKEFFEVANDVIYIYRDEHIYPIYMLAGYHCRNRY QSQGTQSGSALQDPECFALLRFYDQICKEWEGSVPTVLHVGWATFLVQSLGLGTFEGQ VRQKRVTSREAAAEAEPEWGEAREGRRRGRRESKPEPPPPKPKPALDKGLGTGTG GAVSGPPRKPGTVAGTARGPEGSTAOVPAPAAAPPPEGPVLITFOSEKMGKMKELLY ATKINSSAIKQLTQASQVQMKQKVSTFSDYTLSPFLKQRKGL"

175

variation

/gene="MEN1"

/standard_name="L22R"

/note="Lys22Arg"

/replace="g"

357..360

/gene="MEN1"

variation

/standard_name="357del4"

/note="4 bp deletion in kindred 0176"

/replace=""

416

variation

/gene="MEN1"

Db	2342	GGTCCACCTCCAGCCGGGGAGCCTAGGACCACCTAGACCCCTAGGAGTATATTCGCA	2401
QY	8519	CTTCAGAAATTCATATCTTCGGAATCAAGCTCCCTGCCCAATAACTTTCAGTCTGCT	8578
Db	2402	CTTCAGAAATTCATATCTTCGGAATCAAGCTCCCTGCCCAATAACTTTCAGTCTGCT	2461
QY	8579	TCAGAAATTTGGAAATCTAGTTTCCTCTCTTCGTAATCCCGAGTCTGGGACACAAACT	8638
Db	2462	TCAGAAATTTGGAAATCTAGTTTCCTCTCTTCGTAATCCCGAGTCTGGGACACAAACT	2521
QY	8639	CCGCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTCTGACGAACTGCCCCGGA	8698
Db	2522	CCGCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTCTGACGAACTGCCCCGGA	2581
QY	8699	TCAGAGCAGGACCTCCCTCCGACCTCTCGGAACTCCAGAGGTCCAGCCATCTCGG	8758
Db	2582	TCAGAGCAGGACCTCCCTCCGACCTCTCGGAACTCCAGAGGTCCAGCCATCTCGG	2641
QY	8759	AGCATCCCGGAGGAATCTCGAGAGGGTTAGGAGTGGGTGACAAGAGCCTGATCTTC	8818
Db	2642	AGCATCCCGGAGGAATCTCGAGAGGGTTAGGAGTGGGTGACAAGAGCCTGATCTTC	2701
QY	8819	CTGTTTGTACATAGATTTATTTTCAGTTCCTCCAGAAAGATGAATATTTGTTAAAA	8878
Db	2702	CTGTTTGTACATAGATTTATTTTCAGTTCCTCCAGAAAGATGAATATTTGTTAAAA	2761
QY	8879	AAATATAAA 8887	
Db	2762	AAAAAAA 2770	
RESULT 12			
LOCUS	BC002664	1691 bp mRNA PRI 16-MAR-2001	
DEFINITION	Homo sapiens, clone IMAGE:3605739, mRNA, partial cds.		
ACCESSION	BC002664		
VERSION	BC002664.1 GI:12803658		
KEYWORDS	human.		
SOURCE	human.		
ORGANISM	Homo sapiens		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
TITLE	1 (bases 1 to 1691)		
JOURNAL	Direct Submission		
REMARK	Submitted (05-FEB-2001) National Institutes of Health, Mammalian		
COMMENT	Gene Collection (MGC), Cancer Genomics Office, National Cancer		
	Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		
	USA		
	NIH-MGC Project URL: http://mgc.nci.nih.gov		
	Contact: Robert Strausberg, Ph.D.		
	Tel.: (301) 496-1550		
	Email: Robert_Strausberg@nih.gov		
	Tissue Procurement: ATCC		
	cDNA Library Preparation: Rubin Laboratory		
	DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	Sequencing by: National Institutes of Health Intramural		
	Sequencing Center (NISC),		
	Gaithersburg, Maryland;		
	Web site: http://www.nisc.nih.gov/		
	Contact: nisc.mgc@nih.gov		
	Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,		
	Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,		
	Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,		
	Lim, M., Maduro, O.L., Mastello, C., Mastrian, S.D., McCloskey, J.C.,		
	McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,		
	Tiongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,		
	Zhang, L.-H. and Green, E.D.		
	Clone distribution: MGC clone distribution information can be found		
	through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov		
	Series: IRAL Plate: 12 Row: 1 Column: 3.		
	Location/Qualifiers		
FEATURES			

source	1..1691		
	/organism="Homo sapiens"		
	/db_xref="taxon:9606"		
	/clone="IMAGE:3605739"		
	/tissue_type="Uterus, endometrium adenocarcinoma"		
	/clone_lib="NIH_MGC_44"		
	/lab_host="DH10B-R"		
	/note="Vector: pOTB7"		
CDS	<1..861		
	/codon_start=1		
	/product="Unknown (protein for IMAGE:3605739)"		
	/protein_id="AAH02864.1"		
	/db_xref="GI:12803659"		
	/translation="GTRHCRNRNRREALQAWADTATVIQDYNRYREDEEYKEFFEVA		
	NDVPLNLKEAASLLLEAGEBERGEGESQSGSALQDPECFALLRLFYDGIKWEFG		
	STPVLHVGMATFLVQSLGRFEGOVQKRVIVSREAEAEAEAEAEAEAEAEAEAEAE		
	RSKPPEPPPPKPKPALDKGLGTGCGAVSGPPPKPGTAVGTARGPPEGGSTAQVPAPAA		
	SPPEGPVLTFQSEKMKMKELLVATRKINSSAIKLQLTAAQSQVMKKQKVSTPDSYTL		
	SFLKRRKGL"		
BASE COUNT	392 a 540 c 463 g 296 t		
ORIGIN			
	Query Match 13.7%; Score 1253.2; DB 91; Length 1691;		
	Best Local Similarity 96.2%; Pred. No. 1.1e-255;		
	Matches 1337; Conservative 0; Mismatches 48; Indels 5; Gaps 5;		
QY	7499	GGATGCCAGAGCAGGTCCTGGAGTTCAGCCACTGGCGGCAACCTTCTCTCACCTT	7558
Db	300	GGAGGCGAGTCCCAACGCTGTGTGCACGTGGGTGGGCCACCTTTCTTGCAGTCCCT	359
QY	7559	GCTCTCCCACTGCCCCAGGTGGCGGAGAGGTGCGCATAGTACAGCGAGAGCGGAGC	7618
Db	360	AGCGCTTTTGGAGGACAGTGGCGAGAGGTGCGCATAGTACAGCGAGAGCGGAGC	419
QY	7619	GGCGGAGCGGAGAGCGGTGGGGCGAGAGAGCGCGGAGAGCGCGGGGCGGCCACG	7678
Db	420	GGCGGAGCGGAGAGCGGTGGGGCGAGAGAGCGCGGAGAGCGCGGGGCGGCCACG	479
QY	7679	GCAGGAGTCCCAAGCCAGAGAGCGCGCGCCCAAGCAAGCAGCTGGACAAGGCGCT	7738
Db	480	GCAGGAGTCCCAAGCCAGAGAGCGCGCGCCCAAGCAAGCAGCTGGACAAGGCGCT	539
QY	7739	GGGCAACCGCGGAGGTGACGTGTCAGGACCGCGCGCGGAGCGCTCTGGGACTGTG	7798
Db	540	GGGCAACCGCGGAGGTGACGTGTCAGGACCGCGCGCGGAGCGCTCTGGGACTGTG	599
QY	7799	CACAGCCGAGGCGCTGAAGGTGGCGAGCAGCGGTGCGGACCGCGCGGAGCATCAC	7858
Db	600	CACAGCCGAGGCGCTGAAGGTGGCGAGCAGCGGTGCGGACCGCGCGGAGCATCAC	659
QY	7859	ACCGCGGAGGTGCGGAGGTGCTCACTTTCCAGAGTGAGAAGTGAAGGCGAGGAGCT	7918
Db	660	ACCGCGGAGGTGCGGAGGTGCTCACTTTCCAGAGTGAGAAGTGAAGGCGAGGAGCT	719
QY	7919	GCTGGTGGCCACCAAGATCAACTCGAGCGCATCAAGCTGCAACTCAGCGCACAGTCG	7978
Db	720	GCTGGTGGCCACCAAGATCAACTCGAGCGCATCAAGCTGCAACTCAGCGCACAGTCG	779
QY	7979	AGTCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTACACTCTGTCTTCTCAA	8038
Db	780	AGTCAGATGAAGAAGCAGAAAGTGTCCACCCCTAGTGTACTACACTCTGTCTTCTCAA	839
QY	8039	CGCGCAGCGAAGGCGCTCTCACTACTGGGAGTCTGGACCGCTTCTGGGAGCCAGGC	8098
Db	840	CGCGCAGCGAAGGCGCTCTCACTACTGGGAGTCTGGACCGCTTCTGGGAGCCAGGC	899
QY	8099	TCCG-CCTTAGTCTCCCAACTCTGAGCCCTCTGCTGCCCCCAGCCCAAGGCGACAGC	8157
Db	900	TCCG-CCTTAGTCTCCCAACTCTGAGCCCTCTGCTGCCCCCAGCCCAAGGCGACAGC	959
QY	8158	CTCACCTCTACCCCAACCCCTAGGTTCCCGGTCCCGAGTACAGTCTGTATCAACCCACGA	8217

Db 960 CTCACCTTACCAACACCTAGTTCCGGTCCCGAGTACAGTCTGTATCAAAACCCACGA 1019
Qy 8218 TTTTCTCCAGTCAAGACCCAGGGCTCTGCCAGTCTGTAGATATAGTCTCTTCTCC 8277
Db 1020 TTTTCTCCAGTCAAGACCCAGGGCTCTGCCAGTCTGTAGATATAGTCTCTTCTCC 1079
Qy 8278 CAGATCCACGCGGCCAATGGAACCTCAGCTGGTCTCTAATACCAGTCTTTAAAGG 8337
Db 1080 CAGATCCACGCGGCCAATGGAACCTCAGCTGGTCTCTAATACCAGTCTTTAAAGG 1139
Qy 8338 CCCAGCCCTAGAACCCCAAGCTCTCTCGAAGCCGCTACCTAGACCCAGACCAAGT 8397
Db 1140 CCCAGCCCTAGAACCCCAAGCTCTCTCGAAGCCGCTACCTAGACCCAGACCAAGT 1199
Qy 8398 TACTCAGGCTCTCCACAGCTTGTAGGAGCTGAGTCTTACCCCTTAACCCAGGAGCAC 8457
Db 1200 TACTCAGGCTCTCCACAGCTTGTAGGAGCTGAGTCTTACCCCTTAACCCAA -GGAGCAC 1258
Qy 8458 AGGTCCCACTCCAGCCCGGGAGCCTAGGACCACTACGCCCTAGGAGTATATTTCCGC 8517
Db 1259 AGGTCCCACTCCAGCCCGGGAGCCTAGGACCACTACGCCCTAGGAGTATATTTCCGC 1317
Qy 8518 ACTTCAGATTCATATCTTGGGAATCAAGCTCCCTGCCCAATTAATTCAGTCTCTGC 8577
Db 1318 ACTTCAGATTCATATCTTGGGAATCAAGCTCCCTGCCCAATTAATTCAGTCTCTGC 1377
Qy 8578 TTCCAGAAATTTGGAATCCTAGTTTCTCTCTCTCGATCCGAGTCTGGGACACAAAC 8637
Db 1378 -TCCAGAAATTTGGAATCCTAGTTTCTCTCTCTCGATCCGAGTCTGGGACACAAAC 1436
Qy 8638 TCCGCCCCCAGCCTATGAGCATCTTCCAGCCCGCCCTCTTCTGACGAACTGGCCCCCG 8697
Db 1437 TCCGCCCCCAGCCTATGAGCATCTTCCAGCCCGCCCTCTTCTGACGAACTGGCCCCCG 1496
Qy 8698 ATCAGACGAGACCTCCCTCTCGACCTCTGGGAACCTCCAGAGGTCAGCCCATCTCG 8757
Db 1497 ATCAGACGAGACCTCCCTCTCGACCTCTGGGAACCTCCAGAGGTCAGCCCATCTCG 1556
Qy 8758 GAGCATCCCGAGGAAATCTCGCAGAGGGTGTAGGAGTGGGTGACAGGCTGTATCTCT 8817
Db 1557 GAGCATCCCGAGGAAATCTCGCAGAGGGTGTAGGAGTGGGTGACAGGCTGTATCTCT 1615
Qy 8818 CTTGTTTGTACATAGATTTATTTTTCAGTCCCAAGAAAGATGAATATATTTGTTAAA 8877
Db 1616 CTTGTTTGTACATAGATTTATTTTTCAGTCCCAAGAAAGATGAATATATTTGTTAAA 1675
Qy 8878 AATATATAA 8887
Db 1676 AAAAAAAAA 1685

RESULT 13
BC002544
LOCUS BC002544 2633 bp mRNA PRI 16-MAR-2001
DEFINITION Homo sapiens, Similar to multiple endocrine neoplasia I, clone
MGC:1554, mRNA, complete cds.
ACCESSION BC002544
VERSION BC002544.1 GI:12803440
KEYWORDS MGC.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 2633)
TITLE Direct Submission
JOURNAL Strausberg, R.
Submitted (05-FEB-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>
CONTACT: Robert Strausberg, Ph.D.
COMMENT Tel: (301) 496-1550

Email: Robert_Strausberg@nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Shevchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
Lim, M., Maduro, O.L., Masello, C., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
Tiongson, E.E., Touchman, J.W., Tsugeon, C., Vogt, J.L., Walker, M.A.,
Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 4 Row: f Column: 7.

FEATURES

source

1..2633

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="MGC:1554 IMAGE:3138471"

/tissue_type="Placenta, choriocarcinoma"

/clone_lib="NIH_MGC_21"

/lab_host="DH10B-R"

/note="vector: pOTB7"

78..1805

/codon_start=1

/product="Similar to multiple endocrine neoplasia I"

/protein_id="AAH02544.1"

/db_xref="GI:12803441"

/translation="MGLKAAQKTLPLRSIDDDVVRLEAELGREEDPLVLLSILVGFV
EHLAVNRVPTNPVPELTPSPADPDPGLTYPPVADLSITAIARFYARFPAQJRGAVD
LSLYPRGGVSSRELVRKVSIVWNSLSRYSFKDRAHISLFSITGTGLDSSGVAFA
VVGACALGRDLVHLSLSDHAWSLYKGSYMRDRKMEVAFVCAINPISDLHTDS
LELIQLOKLMLLYDLGLHRYPMALGNLADLELEPTGPDPPLLYIKIGIASAKT
VYRDEHYPMYLAGYHCRNRNVRPEOSQTSQGSALQDPCEFAHLLRDKTCWEEGS
DVIPNLKAEASLLLEAGEERPEOSQTSQGSALQDPCEFAHLLRDKTCWEEGS
PTPLVHVWATFLVQSLRGEGQVRQVRYVREAAEAEEPEHGEARERRGRPRR
ESKPEPPPPKPKALDGLGTGAGVSGPPRKPPTGTAGTARGPEGEAQVAPPAAS
PPPEPVLTFTQSEKMKMKELLVATKINSSAIKLQLTQAQSOVMKKOKVSTPSDTLS
FLKRRKGL"

CDS

BASE COUNT 548 a 865 c 729 g 491 t

ORIGIN

Query Match 13.6%; Score 1251.6; DB 91; Length 2633;
Best Local Similarity 96.1%; Pred. No. 2.3e-255;
Matches 1336; Conservative 0; Mismatches 49; Indels 5; Gaps 5;

Qy 7499 GGATGGCCAGACGAGGCTCTGGAGTTCAGCCACTGGCCGCAACCTTGTCTCACCTT 7558

Db 1244 GAGGAGGAGTCCAGCCGCTGTGTGCGGCTGGGCTTCTTGTGTCAGTCCCT 1303

Qy 7559 GCTCTCCCACTGGCCGAGGTCGCGCAGAGGTGCGCATAGTGAGCCGAGAGGCTGAGCG 7618

Db 1304 AGGCCGTTTGGAGGACAGGTGCGGACAGAGGTGCGCATAGTGAGCCGAGAGGCTGAGCG 1363

Qy 7619 GGCCGAGCCGAGGAGCGTGGGGCGAGAGCCCGGGAAGCCGCGCGGGGGGGCCACG 7678

Db 1364 GGCCGAGCCGAGGAGCGTGGGGCGAGAGCCCGGGAAGCCGCGCGGGGGGGCCACG 1423

Qy 7679 GCGGAGTCCAGGACGAGGAGCCCGCCCGCCCAAGAGCCAGCACCTGGACAGGGCT 7738

Db 1424 GCGGAGTCCAGGACGAGGAGCCCGCCCGCCCAAGAGCCAGCACCTGGACAGGGCT 1483

Qy 7739 GGGCACCGCCAGGCTGAGTGTCTAGGAGCCCGCCCGGGAAGCCCTCTGGGACTGTCCCTGG 7798

Db 1484 GGGCACCGCCAGGCTGAGTGTCTAGGAGCCCGCCCGGGAAGCCCTCTGGGACTGTCCCTGG 1543


```
Qy 7799 CACAGCCGAGCCCTGAAGGTGGCAGCAGCGCTCAGTGTCCAGCAGCCGCGAGCATCAC 7858
Db 766 CACAGCCGAGCCCTGAAGGTGGCAGCAGCGCTCAGTGTCCAGCAGCCGCGAGCATCAC 825
Qy 7859 ACCGCCGAGGCTCAGTGTCTACTTCCAGAGTGAAGATGAAGGCGCATGAAGGAGCT 7918
Db 826 ACCGCCGAGGCTCAGTGTCTACTTCCAGAGTGAAGATGAAGGCGCATGAAGGAGCT 885
Qy 7919 GCTGTGTGCCACCAAGATCAACTCGAGCGCATCAAGCTGCAACTCAGGCGACAGTGC 7978
Db 886 GCTGTGTGCCACCAAGATCAACTCGAGCGCATCAAGCTGCAACTCAGGCGACAGTGC 945
Qy 7979 AGTCAGATGAAGAGCAGAGAGTGTCCACCCCTAGTGAAGTACACTCTGTCTTCTCAA 8038
Db 946 AGTCAGATGAAGAGCAGAGAGTGTCCACCCCTAGTGAAGTACACTCTGTCTTCTCAA 1005
Qy 8039 CGCGCAGGCGAAGCCCTCAACTACTGCGGACTTGGGACCGTGTGGGACCCAGGC 8098
Db 1006 CGCGCAGGCGAAGCCCTCAACTACTGCGGACTTGGGACCGTGTGGGACCCAGGC 1065
Qy 8099 TCCG-CCTTAGTCCCCCAACTCTGAGCCCATGTTCTGCCCCCAGCCAAAGGGACAGGC 8157
Db 1066 TCCGCCCTTAGTCCCCCAACTCTGAGCCCATGTTCTGCCCCCAGCCAAAGGGACAGGC 1125
Qy 8158 CTCACCTCTACCCAAACCTAGTTCGCGGTCCGAGTACAGTCTGTATCAAAACCCACGA 8217
Db 1126 CTCACCTCTACCCAAACCTAGTTCGCGGTCCGAGTACAGTCTGTATCAAAACCCACGA 1185
Qy 8218 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCAGTGGTTAGATATAGTCTCTCTCC 8277
Db 1186 TTTTCTCCAGCTCAGAACCCAGGCTCTGCCAGTGGTTAGATATAGTCTCTCTCC 1245
Qy 8278 CAGATCCACGCGCCCAATGGAACCTCAGCGTGGGTCTTAATACAGCTTTTAAAGG 8337
Db 1246 CAGATCCACGCGCCCAATGGAACCTCAGCGTGGGTCTTAATACAGCTTTTAAAGG 1305
Qy 8338 CCCAGCCCTAGAAACCCAGCTCTCTCGGACCGCTACCTAGAGCCAGACCAAGT 8397
Db 1306 CCCAGCCCTAGAAACCCAGCTCTCTCGGACCGCTACCTAGAGCCAGACCAAGT 1365
Qy 8398 TACTCAGGCTCTCCAGCTTGTAGAGCTGAGTGTTCACCCCTTAACCCAGGAGCAC 8457
Db 1366 TACTCAGGCTCTCCAGCTTGTAGAGCTGAGTGTTCACCCCTTAACCCAA-GGAGCAC 1424
Qy 8458 AGGTCCCACTCCAGCCGCGGAGCCCTAGGACCACTCAGCCCTTAGGATATATTCCGC 8517
Db 1425 AGGTCCCACTCCAGCCGCGGAGCCCTAGGACCACTCAGCCCTTAGGATATATTCCGC 1483
Qy 8518 ACTTCAGAAATTCATATCTCGGAATCCAAAGCTCCCTGCCCAATACCTCAGTCTCTGC 8577
Db 1484 ACTTCAGAAATTCATATCTCGGAATCCAAAGCTCCCTGCCCAATACCTCAGTCTCTGC 1543
Qy 8578 TTCAGAAATTTGGAATCTAGTTCTCTCTCTCGTATCCGAGTCTGGGACACAAAC 8637
Db 1544 -TCCAGAAATTTGGAATCTAGTTCTCTCTCTCGTATCCGAGTCTGGGACACAAAC 1602
Qy 8638 TCCGCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTTCCAGCAAACTGCGCCCGG 8697
Db 1603 TCCGCCCCAGCCTATGAGCATCTGAGCCCGCCCTCTTCTTCCAGCAAACTGCGCCCGG 1662
Qy 8698 ATCAGAGCAGACCTCTCTCGACCTCTCGGACCTCCAGAGGTCACCCATCTCG 8757
Db 1663 ATCAGAGCAGACCTCTCTCGACCTCTCGGACCTCCAGAGGTCACCCATCTCG 1722
Qy 8758 GAGCATCCCGAGGAATCTCCAGAGGGTTAGGAGTGGGTGACAAGAGCTGATCTCT 8817
Db 1723 GAGCATCCCGAGGAATCTCCAGAGGGTTAGGAGTGGGTGACAAGAGCTGATCTCT 1781
Qy 8818 CTTGTTTTGATAGATTTATTTTTCAGTTTCCAGAAAGATGAATACATTTTGT 8873
Db 1782 CTTGTTTTGATAGATTTATTTTTCAGTTTCCAGAAAGATGAATACATTTTGT 1837
```

```
RESULT 15
AF109390 LOCUS
DEFINITION Mus musculus menin (Men1) gene, complete cds.
ACCESSION AF109390
VERSION AF109390.1 GI:5051649
KEYWORDS house mouse.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 9286)
Guru,S.C., Crabtree,J.S., Brown,K.D., Dunn,K.J., Manickam,P.,
Prasad,N.B., Wangsa,D., Burns,A.L., Spiegel,A.M., Marx,S.J.,
Pavan,W.J., Collins,F.S. and Chandrasekharappa,S.C.
Isolation, genomic organization, and expression analysis of men1,
the murine homolog of the MEN1 gene
Mamm. Genome 10 (6), 592-596 (1999)
99272818
PUBMED 10341092
2 (bases 1 to 9286)
Guru,S.C., Crabtree,J.S., Brown,K.D., Dunn,K.J., Manickam,P.,
Prasad,N.B., Wangsa,D., Burns,A.L., Spiegel,A.M., Marx,S.J.,
Pavan,W.J., Collins,F.S. and Chandrasekharappa,S.C.
Direct Submission
Submitted (25-NOV-1998) Genetics and Molecular Biology Branch,
National Human Genome Research Institute, NIH, Bldg49, Rm3C36, 9000
Rockville Pike, Bethesda, MD 20892, USA
FEATURES
Location/Qualifiers
1..9286
/organism="Mus musculus"
/strain="129Sv/c37"
/db_xref="taxon:10090"
/chromosome="19"
2962..8815
/gene="Men1"
join(2963..3829,4649..4857,5077..5205,5403..5443,
5525..5612,6207..6343,6750..6885,7263..7427,7611..8096)
/gene="Men1"
/product="menin"
join(2963..3006,3369..3829,4649..4857,5077..5205,
5403..5443,5525..5612,6207..6343,6750..6885,7263..7427,
7611..8096)
/gene="Men1"
/product="menin"
join(3385..3829,4649..4857,5077..5205,5403..5443,
5525..5612,6207..6343,6750..6885,7263..7427,7611..8096)
/gene="Men1"
/note="similar to Homo sapiens MEN1"
/codon_start=1
/product="menin"
/db_xref="GI:5051650"
/db_xref="AF109390"
/translation="MGLKAAQKTLPLRSTDDVVRFLFAALGREEDPLVLSLVGVF
EHLFANVRVPTWVPELTQFSPADPPGGLTYFPVADLSIIAALYARFTQIRGAVD
LSLPRGGVSSRELKVKSDVWNSYFKDRAHQISLFSFTITKLDSSGVAER
VGCAGALGRDLVHLALSDHAWVFGNGEOTAETVHGKNGDRGQTVNAGVAER
SWLYLKSGYMRCKMEVAFVCAINPSIDLHTDSLELLQLOKLLWLLYDLGLHRY
PMALGNLADLELEPTGPRDPLTLVHKGASAKTYVODEHIYPMYLAGYHCRNVV
REALQAWADTATVIQDYNCREDEELYKEFEVANDVIPNLKEASLLETGERTGE
QAGTCGQSALQDPECFAHLLRFYDGIKKBEGSPPTVLHVGNATFLVSLGRFEQ
VRQKHIVISREAEAAEPEWDEARERRRGRPRRESKPEEPPTPKPALDKGPGSQ
SAGSGPRKTSQVPTGTGGVQNAQAAPAAAPAPPEPGLVITFQSEKMKMKKELL
VATKINSSAIKLLQTAQSQVMKKQKVTSPDVTLSFLKQRKGL"
BASE COUNT 2104 a 2383 c 2518 g 2281 t
ORIGIN
```

```
Query Match 11.7%; Score 1075.8; DB 94; Length 9286;
Best Local Similarity 60.1%; Pred. No. 3,7e-218;
Matches 2957; Conservative 0; Mismatches 1473; Indels 488; Gaps 51;
Qy 4272 CTTTCATTACCTCCCTTCCACAGGACCAAAATGGACACTCCGGTCTCGCTTGTCT 4331
```

Db 4624 CTCTATGACTGCGCTTTTACCACAGGCACCAAACTGGACAGCTCGGGCGTGGCTTTGCA 4683
QY 4332 GTGGTTGGGGCCCTCCAGGCGCTTCCGGGTCTCCGGGTATGTCACCTCGCCCTGTCTGAGGAT 4391
Db 4684 GTGGTAGGGCCCTCCAGGCTCTGGGTCTCAGAGATGTCATCTGGCCCTGTCTGAAGAT 4743
QY 4392 CATGCCCTGGGTAGTGTGTTGGGCCCAATGCGGAGCAGACAGCTGAGGTACCTGGCAGGC 4451
Db 4744 CATGCTTGGGTGGTGTGTTGGGCCCAACGGGGAGCAGACAGCTGAGGTGACGTGGCAGGC 4803
QY 4452 AAGGCAACGAGGACCGCAGGGGCCAGACAGTCAATGCCGGTGTGCTGACGGGTATTG 4511
Db 4804 AAAGGCAACGAGGACCGCAGAGCCAGACAGTCAATGCCGGTGTGCTGACGGGTATTG 4863
QY 4512 TTCCCTCCCCCAGCGCTTGTCCCTTTCATATGTAAGTCAAGCCACCAAGGACTC 4571
Db 4864 TATCTTCCCTCCA---TTGTGCCCTTTTCCCTTCTGTGAGTCTAAGTGACCTGAGGATTC 4920
QY 4572 CATTTTCTTGGGCCACAC-----CCCTTTCTTCCCATACCACCCACATAGG 4618
Db 4921 TGCTTCCCGCAGCGCTCCCTTACCAGTCTCCCTCTGTGTTTATGTATCCAAAGCGGGG 4980
QY 4619 AAGGGAACAGAGAGCCCTTTTCTGCTGTCTCATCTCCCTGAAGCAGGACAGGCTGG 4678
Db 4981 GCGAGGCGAGGCGAGGTGAATTGATTTACGTTTGTGCTGTGCCCTAGAGCAGGCTCAAGGTG 5040
QY 4679 GCATCATGAGACATATGATGATCATCCCCCTTAGAGCTGGCTGTACCTGAAAGGATC 4738
Db 5041 TACATCTGAGATGT-CTGCTTTTCTCCCTTCCAGAGCTGGCTGTACCTGAAAGGGTC 5099
QY 4739 ATATGCGCTGTGACCGCAAGATGAGGTGGGTTTCATGGTGTGCCATCAACCCCTTC 4798
Db 5100 GTACATGCGCTGCGACCGTAAGATGAGGTGGGTTTCATGCTGTGTCCTATCAACCCCTTC 5159
QY 4799 CATTGACCTGCACACCGACTCGCTGAGCTCTTGACAGCTCCAGCAGGTGAGGCTGAGCC 4858
Db 5160 CATCGATCTTACACTGACTCTTTGGAACCTTTGCAACTACAGCAGGTGAGGACTGCC 5219
QY 4859 AATGGGCGAGGACTGGGCTAGGCCAGACTTGACTTGCTGTGGACCCCTGGGCGAGGSCAC 4918
Db 5220 -----AATGCACACGAAAG 5219
QY 4919 TTTCCCTTCTGAGCTTACGTTTCCCTCTCGGAAATGGGTAGTAATTTCTTGGGCTG 4978
Db 5220 ---CCAGCCCGAGGACCTTCCCTTCCCTTCTGTACAGAGAACTGATCCTTTTC- 5269
QY 4979 GCCTTTCCAGGCTCTTGGGAGTAGAATTGAGATGTGAATGCTTTGACTCCATTA 5038
Db 5270 -----AATGCACACGAAAG 5283
QY 5039 AAGGCTGGTCCAGAAATTTTGGCCCTTCCACATGGTGGGTGCTGTTGTTCTGAC 5098
Db 5284 GCCTTTGAGTCTTTACTCTAGCCCTTC---TGTGGTGGAGGGCGCTGTTTCTCTGCT 5341
QY 5099 CCCCACCTCTGCCCGATGAGCTAAGGACCGTTCTCTCCCTGTTCGGTGGCTCATAACT 5158
Db 5342 CACTTCTAGCTAGCTGGG-----GCTCCCTTCTACTTTACCAACGCC 5384
QY 5159 CTCTCTTCCGCTCTAGAGCTGCTCTGCTGCTATGACCTGGGACATCTGGAAGG 5218
Db 5385 CTCTCCCTTTGGCTCTAGAGCTGCTCTGCTGCTGTATGACCTCGGACATCTGGAAGG 5444
QY 5219 TCAGTAGAGGAAGTGGCCAGGCTGCGCTGTGAGCGCGGGGCTGGTGGCAGCCTG 5278
Db 5445 TCAGTGAAGGCTATGATCAGATAGGCTTTGTGGGATCAGGAACTGAGTGATAGCCTG 5504
QY 5279 A-ATTATGATCTTTCTTAGTACCCCATGGCTTAGGAACTCTAGGAGCTCTAGAGGAG 5337
Db 5505 ATTTTATGATCTGTCCAGATACCCCATGGCTAGGAACTTGGCAGACCTTGGAGGAG 5564
QY 5338 CTGAGCCCAACCTGCGCGGCGAGCCCACTACCTCTTACCAAGAGGTGGGGGCACT 5397
Db 5397

Db 5565 CTGAGGCTTACCCCGCGGCCAGACCCCACTCACCTTATCACAAAGGTAAGGGCATCT 5624
QY 5398 AAGGAGGCTCAGAAGGAGAGACCTTAACAGTGGCTAGGCGAGGGCCCTCATCTGGGCAG 5457
Db 5625 CA-----GAAACAGCAGCTCTTCTCTGGGATGAGAGGAG-----AGAGGTG 5668
QY 5458 ATGAGAAGAGAACTTTGTGTGTTGGGGGTATGCCCATCCAGTCTCACCTTGTGTGTAAC 5517
Db 5669 AAAAGGGAAGAACTCTGTGTGTTGGAGGTAT-----TTCCCTGCTTACTTTGTGTAGGG 5723
QY 5518 TGTGTGCAGAACTCAGTTTCACTCAGGCTGTCTGAGGGGTCTCCAGGGTTCGCCAGCCTGG 5577
Db 5724 CATGTGTGAATCAGGGTACTTCTGTGTTTAAAGAG-----CCCTGGCTCCCCAGTAG 5780
QY 5578 GAGTGGCAGGGCTGCAATTTGTCCCTCAGCCCTGCTTTTCTGCCACTGCTTACTCTCC 5637
Db 5781 GAGTGGCAGGGCCGATTTGATCCCAAGCCCTGCAGTTTTCAGCTGCCCTTTATTATCC 5840
QY 5638 TTCTTGAGTATTAACAGAGTCAAAATGTGTAGAGCAGCTGAAGAGGGGTGTTCACATTG 5697
Db 5841 TGGCTATAGCTCTGGTAG---CACTGTGGAGAGATGTTAGCTTGTCTCATTTGTGGTAGGA 5897
QY 5698 GTGGGTAGTGGGAGGAGGCGCATTTGGCTGGGCTTGAAGTCTTTTGTGTATGTA 5757
Db 5898 GGCTGGTTGGGCTTGAAGTCTTTTACACAGGACCATGTGGTGAAGAAACCTGCCCCAGGCT 5957
QY 5758 GAAGAGTGTCTGAGAAAGAGAGGCGCTCAGCTCGGAGGGCAGGCGCCCTCGAGT 5817
Db 5958 TGAGTGTTCCTGCTCCTCCTAGGCTTGTGTTTGTGGGT---CACCACTCTCTACTGT 6014
QY 5818 CTCCCGAGGCTCAGCAGCAGTCTGTAGACCCAGGGAGGAGACCAAGGTAGAAAGGCT 5877
Db 6015 CTG-CCCTTCTTTAGCAAGCAGGCTTGTAGATCCAGGCGCTGGAATGAGGTGGAGGGA 6073
QY 5878 GGACGAGTGGAGTGGGAGTGGAGATGGAGAGGACTCCCTGGGATCTTCTCTGCGCC 5937
Db 6074 GTATGGAACACCAAGG-CAAGAGGGGCCAGCAGTATTTAGAGACAGGGAGGACCTTTAGCAT 6132
QY 5938 CTCTCTGGCTGCGCTTGGTGGGCAATTTGTGCCAGCAGGCGAGCTGGGGCTGCCCTCCCTG 5997
Db 6133 CCTTTTCTGT-----AGGCTATCAGTGTATGAGGAGCTGGGGGCTGCCCTCCCTG 6182
QY 5998 AGATCTCTGCTCAGCTCCATCCAGGCAATGCTCTCAGCCAGAGCTACTATCGGGAT 6057
Db 6183 AGGTCTCTCTG---CTCCATCTCCAGGAAATGCCCTCAGCTAAGACCTACTACCAAGAT 6239
QY 6058 GAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTGTCCCAACCCCAATGTGCGG 6117
Db 6240 GAACACATCTACCCCTACATGTACCTGGCTGGCTACCACTGTCCCAACCCCAATGTGCGG 6299
QY 6118 GAAGCCCTCAGGCTGGGCGGACACGGCCACTGTCTATCCAGGAGTGAAGTCCCCCTAC 6177
Db 6300 GAAGCCCTCAGGCTGGGCGGACACTGCCACTGTTATCCAGAGTGAAGATTCCTCTGC 6359
QY 6178 TAGGGCTCAGGCTGTCTTCTTCCCTCCATCAGTATTTCCAAACCACTCTGCTCCAGGA 6237
Db 6360 CAGGG-TTCCAGGCTGCCCTTGT---CTCTATGACTTTTCCAA--TCGGTCACTCCAGTT 6412
QY 6238 CTGAGGCTGGCTCCCAAGCCCATCCCTTCCATCCAGTCCCTAGGCAGCAAGGCCAC 6297
Db 6413 GTGAGACTTGTCTCAGCTCTGTCTCTCTGATCAGGCCCTTATACAG---TCAAGGCTAC 6469
QY 6298 CATTAACAGGAGTGAAGGACCCCTGATTAAGGTGTACATCTTTCCCTGCCCTCCCTCTC 6357
Db 6470 CATTACTTGGAACTGAGAACCCCTGGTGTAGCTGCTCTCTCTCATGCTTCTGTCA 6529
QY 6358 CTCTAAATTTTCTTCTCAGAAAGTCTCAAATCTCCAAATGTTTAAACCACCACTATCC 6417
Db 6530 CTCTCTCTGCCAGTTTCTTCTCTTATAGTCAGTTGTGTTATCCATGGTGTCTGT 6589
QY 6418 AGCAGTGGGACTTCCACCTCGGGCCCATGCCCTCTCTCTCTCTCTCTCTCTCTCTCT 6477
Db 6590 CTCTGCCCAACTTCTAGTCCAGTCCCACTCCCACTCCCACTTCAACCTTGGCTTGTAGCAG 6649


```

Db 8594 ACATGAAGCTCCGACTTTATCTGTGAATACTGAGCCTCGCCCTTCTTGACCAACTG 8653
Qy 8691 GCGCCGATCAGAGAGGACCTCCCTCCGACCTCTGGGAACCTCCAGAGTCCAGCC 8750
Db 8654 GCGCAGAACAGAGAGGACCT--CTTCCAGCCTCTGGGAACCTCCCTGAGGTC----- 8706
Qy 8751 CATCTCGGAGCATCCGGAGGAAATCTGCAGAGGGTTAGGAGTGGGTGACAAGAGCCTG 8810
Db 8707 -----CCAGGAGGAATCTGCACGGGTGGGGAG-----GGCCA 8741
Qy 8811 ATCTCTTCGTTTGTACATAGATTTATTTTCAGTTTCCAAAGAAAGATGAATACATTTT 8870
Db 8742 GCCTCTTCTGTTTGTACATAGATTTATTTTCAGTTTCCAAAGAAAGATGAATGCAATTT 8801
Qy 8871 GTTAAAAAAATATAAGCGCAAGTCCATGTTTATCTGGGAAATTGGGGATGGGGCGGG 8930
Db 8802 GTTAAGAAAAAGAGATGATGTATGTTTTCATCTGAGAAATCGGGATGGGGCGGG 8861
Qy 8931 AGTGGAGCGCCCTTCTTCCCTTTGCTTCTGCTCCCGGACTTTGCGCTCCCTACCTG 8990
Db 8862 AGTGA-----TATCTGCTTTTGTCTTTTAGGATTTTGAATTTATGCTCTGCACCTG 8915
Qy 8991 TGGAGCGGAGCAGACAGTGGCGGGAAGGACGTAGGCTCCGCCCGGCTTGGGGCTTC 9050
Db 8916 TAGCGG-----GCGAGAATTGGGAGGGGTACCTAAGCTCCGCCCTGCTTCTGGGACTT 8970
Qy 9051 CCGCGCGCGCGCGGCGCCGCTCCCGGGGCGCTCTCCCGGACTGGCGGTGGGGCATC 9110
Db 8971 CCGCGCGCAACCA-----GGGCGGCTCTCTCGGCCAGGCGGGCGGGGG 9016
Qy 9111 CCNGGCGCGCGCGCGCGCGGCTTACGCCCGCGCGCGGCTTACAGAGCCACGG 9168
Db 9017 CAGCCCGGGCTGGCCCGCGCGGCTGCGCCGCCCTACGGCCTCCGAGTCTCGG 9074
```

THIS PAGE BLANK (USPTO)